(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2002 (27.12.2002)

PCT

(10) International Publication Number
WO 02/103363 A2

(51) International Patent Classification[7]: G01N 33/68, C12N 15/00

(21) International Application Number: PCT/GB02/02839

(22) International Filing Date: 18 June 2002 (18.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0114856.8 18 June 2001 (18.06.2001) GB

(71) Applicant *(for all designated States except US)*: **MEDICAL RESEARCH COUNCIL** [GB/GB]; 20 Park Crescent, London W1B 1AL (GB).

(71) Applicants and
(72) Inventors *(for all designated States except US)*: **DE WILDT, Ruud** [GB/GB]; Domantis Limited, Granta Park, Abington, Cambridge CB1 6GS (GB). **TOMLINSON, Ian** [GB/GB]; Domantis Limited, Granta Park, Abington, Cambridge CB1 6GS (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **HOLLIGER, Philipp** [GB/GB]; MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH (GB).

(74) Agents: **MASCHIO, Antonio** et al.; D Young & Co, 21 New Fetter Lane, London EC4A 1DA (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/103363 A2

(54) Title: SELECTION BY AVIDITY CAPTURE

(57) Abstract: A method is provided for selecting first and second molecules according to an interaction between said first and second molecules, the method comprising: (i) providing said first and second molecules; (ii) allowing said first and second molecules in solution, such that a complex comprising each of said first and second molecules in multivalent form is generated; (iii) isolating the complex and characterising the constituent molecules.

SELECTION BY AVIDITY CAPTURE

Field of the invention

The present invention relates to an improved method for identifying molecular interactions.

Background of the invention

The flood of genetic information from whole genome sequencing has resulted in the identification of a large number of genes whose function is unknown or only partially characterised. An important aspect of determining gene function is to identify other gene products or molecules with which a given gene or gene product interacts. To do this effectively requires the use of efficient high-throughput strategies.

For the investigation of protein-protein interactions, two-hybrid technologies pioneered by Fields & Song (1989) have been particularly successful and have spawned a range of related technologies for the investigation of protein-protein, protein-nucleic acid (1- and 3-hybrid) and protein-drug interactions. All of these require a eukaryotic host and consequently molecular repertoire sizes are limited (ca. 10^6).

Larger repertoire sizes are accessible in bacteria and three bacterial two-hybrid systems have been described recently (Karimova, *et al.* (1998) *Proc. Natl. Acad. Sci.USA* 95, 5752-5756; Pelletier, *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95, 12141-12146; Dove, *et al.* (1997) *Nature* 386, 627-630). Two are based on the reconstitution of enzyme function by association of its constituent parts through interaction of proteins fused to these parts, another on transcription activation by increasing affinity of the RNA polymerase for a promoter by protein-protein interaction. One of the split-enzyme approaches has been applied successfully to the isolation of interacting leucine-zipper pairs from a 10^6 starting repertoire (Pelletier *et al.*, 1999, Nature Biotechnology 17(7): 683-90). Neither, however, has so far been applied in a generic way, e.g. the screening of interactions of a given bait protein with members of a cDNA library. Furthermore, all are *in vivo* methods employing either metabolic or antibiotic selection. Because bacterial cells develop a mutator

phenotype under adaptive stress (see for example Sniegowski et al 1997, Nature 387(6634): 703-5) , the generation of false positives by mutation or recombination may limit accessible library size in these systems. Indeed, Karimova, 1998, describe a model selection, where correct positives could not be retrieved from higher dilutions than $1/10^6$.

Phage display has allowed the isolation of many novel peptide and protein ligands and in particular for the isolation of antibodies directly from diverse libraries of V-genes (Winter, *et al.* (1994) *Annu.Rev. Immunol.* 12, 433-455; Winter, 1998; FEBS Letters 430: 92-94). Electroporation into bacteria allows library sizes of up to 10^{11} and using combinatorial infection together with methods of recombination (Waterhouse, 1993 Nucleic Acids Research 21(9): 2265-6; Griffiths, *et al.* (1994) Embo *J,* 13, 3245-60; Fisch, 1996, Proceedings Of The National Academy Of Sciences Of The United States Of America 93(15): 7761-6 Sblattero 2000 Nature Biotechnology 18(1): 75-80), library sizes of 10^{13} or higher are possible.

However, current phage technology is mostly used in the isolation of ligands to a single target. When selecting against complex targets (such as whole cells, cell lysates, membrane preparations, cDNA libraries) strong biases may arise due to differences in phage growth, abundance of target proteins and the number and affinity/binding kinetics of specific binders present in the phage repertoire. Such biases also occur when selecting against single targets and often lead to the isolation of only a portion of the antigen binders present in the repertoire, i.e. the binders which not only bind antigen well but also express and fold well on phage, are non-toxic to the cell, bind the most exposed epitope on the target protein.

Possible solutions are to limit the number of rounds of selection or directly screen unselected libraries (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994). Another approach would be the use of recombinant (and preferably normalised) libraries of expressed cDNAs as a selection target. It would thus be desirable to be able to "cross" phage display libraries with such cDNA expression libraries for the simultaneous isolation of interacting pairs from the two starting libraries comprising for example antibodies and

antigens or generally interacting polypeptides. Ideally, the benefits of the two approaches would be combined.

Within phage display several approaches have been described that could potentially be used for library vs. library selections. These link the protein-protein interaction to phage infectivity. In the SAP (Duenas, 1994, BioTechnology 12(10): 999-1002) or SIP (Krebber, 1995, Febs Letters 377(2): 227-31; 1997, Molecular Biology 268(3): 607-18) approach, the g3p infectivity protein is split with protein-protein interaction reconstituting a link between the two fragments and thus infectivity. In another approach (CLAP), interactions between protein displayed on phage and on the F-pilus of the bacterial cell are required for infectivity (Malmborg, 1997, J.Mol.Biol. 273: 544-551). Both approaches suffer from the fact that even high-affinity protein-protein interactions appear not to reconstitute full phage infectivity and reconstitution may depend on a fairly narrow window of affinities. Furthermore, it has recently been suggested that some formats of SAP/SIP may suffer background infectivity from interactions between the g3p domains (Chatellier, 1999, Febs Letters 463(3): 371-4).

Accordingly, there is a need for an improved method for identifying molecules that interact with one another.

Summary of the invention

Intermolecular interactions occur within a wide range of affinities, ranging from relatively low affinities (10^{-6} M typical of intracellular signalling interactions (e.g. SH3 domains), to medium affinities often found in, for example, antibody-antigen interactions (10^{-8} to10^{-9} M) to very high affinities (10^{-10} to 10^{-12} M), like those, for example, found in growth-factor receptor interactions and some highly affinity matured antibody-antigen interactions. However, especially with low to medium affinity interactions, the half-life of the interactions is often too short compared with the time required to detect the interactions in systems such as phage display.

Thus one aspect of the present invention is based on increasing the apparent affinity (and consequently half-life) of intermolecular interactions by allowing multivalent interactions

to occur. The resulting apparent increase in affinity gained by multivalent interaction, also termed the avidity (or chelate) effect, is due to the fact that when two or more binding interactions take place within the same molecular complex, there is only a very small additional entropic price to be paid for the second or further interactions. This is because most degrees of freedom have already been lost when binding through one binding site immobilised the multivalent ligand. Avidity can have particularly drastic effects on the dissociation kinetics (k_{off}), as all interactions must be broken before dissociation can take place. This effect is observable in the much-increased affinities of multivalent antibody species for solid-phase antigen.

Thus, the invention provides a method for selecting molecules that interact by allowing multivalent interactions to occur between the molecules, thus increasing the stability of any complex formed as compared with a corresponding monovalent interaction.

Accordingly, the present invention provides a method for selecting first and second molecules according to an interaction between said first and second molecules, comprising:

(i) providing said first and second molecules;

(ii) allowing said first and second molecules to interact in solution, such that a complex comprising each of said first and second molecules in multivalent form is generated;

(iii) isolating the complex and characterising the constituent molecules.

Preferably, the first and/or second molecule have been modified to increase their valency. Preferably, one or more reactive groups is present on said first or second molecules, or both, such that first (and/or second) molecules associate with each other to form multivalent first molecule complexes and/or multivalent second molecule complexes. The multivalent complexes may be selected for binding according to the invention.

In an embodiment of the invention, the reactive groups form a covalent bond on interaction. For example, the groups may be enzyme-suicide substrate pairs. However, covalent bond formation is preferably controlled, in particular when screening interactions, as illegitimate disulphide bond formation may lead to false positives. This may be

controlled by conducting experiments in appropriate buffer conditions, e.g. in the presence of reducing agents such as > 1mM DTT.

In a preferred embodiment, the first and second molecules are polypeptides, more preferably the first and/or second molecules are provided as a plurality of first and second polypeptides. Accordingly, the present invention also provides a method for selecting first and second polypeptides from first and second pluralities of polypeptides according to an interaction between said first and second polypeptides, comprising:

(i) providing a first library of polynucleotides encoding a plurality of first polypeptides and a second library of polynucleotides encoding a plurality of second polypeptides;

(ii) expressing said first and second libraries to generate the first and second pluralities;

(iii) allowing said first and second pluralities to interact in solution, such that a complex comprising each of said first and second polypeptides in multivalent form is generated;

(iv) isolating the complex and characterising the constituent first and second polypeptides.

In a particularly preferred embodiment, the second molecule is multivalent. Thus, the present invention further provides a method for selecting specific interacting pairs comprising a first molecule and a second molecule which method comprises:

(i) contacting a plurality of first molecules with a second molecule, wherein said second molecule is multivalent and wherein the contacting takes place under conditions that permit interaction of the second molecule with one or more of the first molecules to form a multivalent complex comprising a first molecule and a second molecule;

(ii) isolating a complex formed in step (i) and;

(iii) characterising the first molecule in said complex.

The second molecule may be provided as a plurality of second molecules and thus the present invention also provides a method for selecting interacting pairs comprising a first molecule and a second molecule which method comprises:

(i) contacting a plurality of first molecules with a plurality of second molecules, wherein said second molecules are multivalent and wherein the contacting takes place under conditions that permit interaction of one or more of the second molecules with one or more of the first molecules to form a multivalent complex comprising a first molecule and a second molecule;

(ii) isolating a complex formed in step (i) and;

(iii) characterising the constituent first and second molecules in said complex.

In a preferred embodiment of the present invention, where the first and/or second molecules are polypeptides, the polypeptides are provided by expression from polynucleotides encoding said polypeptides. Accordingly, the present invention also provides a method for selecting binding pairs comprising a first polypeptide and a second polypeptide which method comprises:

(i) expressing a plurality of first polynucleotides to produce a plurality of first polypeptides such that each first polynucleotide which directs expression of a corresponding first polypeptide is associated with said first polypeptide

(ii) contacting said plurality of first polypeptides with a second polypeptide, wherein said second polypeptide is multivalent and wherein the contacting takes place under conditions that permit multivalent binding of the second polypeptide to one or more of the first polypeptides.

(iii) selecting one or more first polypeptides that bind to the second polypeptide;

(iv) isolating the corresponding first polynucleotide(s) that encode(s) the one or more first polypeptides selected in step (iii).

In another preferred embodiment of the present invention, where the first and/or second molecules are polypeptides, the polypeptides are provided by expression from polynucleotides encoding said polypeptides. Accordingly, the present invention also provides a method for selecting binding pairs comprising a first polypeptide and a second polypeptide which method comprises:

(i) expressing a plurality of first polynucleotides to produce a plurality of first polypeptides such that each first polynucleotide which directs expression of a corresponding first polypeptide is associated with said first polypeptide, and expressing a plurality of second polynucleotides to produce a plurality of second

polypeptides such that each second polynucleotide which directs expression of a corresponding second polypeptide is associated with said second polypeptide,

(ii) contacting said plurality of first polypeptides with a plurality of second polypeptides, wherein said second polypeptides are multivalent and wherein the contacting takes place under conditions that permit multivalent binding any of the second polypeptides to one or more of the first polypeptides.

(iii) selecting one or more interacting pairs of a first polypeptide that bind to a second polypeptide;

(iv) isolating the corresponding first and second polynucleotide(s) that encode(s) the one or more first and second polypeptides selected in step (iii).

In yet another preferred embodiment of the present invention, where the first and/or second molecules are polypeptides, the polypeptides are provided by expression from polynucleotides encoding said polypeptides. Accordingly, the present invention also provides a method for selecting binding pairs comprising a first polypeptide and a second polypeptide which method comprises:

(i) expressing a plurality of polynucleotides to produce a plurality of first and second polypeptides such that each polynucleotide which directs expression of a corresponding first and second polypeptides is associated with said first polypeptide

(ii) contacting said plurality of first polypeptides with the plurality of second polypeptides expressed from the polynucleotide, wherein said second polypeptides are multivalent and wherein the contacting takes place under conditions that permit multivalent binding any of the second polypeptides to one or more of the first polypeptides.

(iii) selecting one or more interacting pairs of a first polypeptide that bind to a second polypeptide;

(iv) isolating the corresponding polynucleotide(s) that encode(s) the one or more first and second polypeptides selected in step (iii).

In another preferred embodiment of the present invention, where the first and/or second molecules are polypeptides, the polypeptides are provided by expression from polynucleotides encoding said polypeptides. Accordingly, the present invention also provides a method for selecting binding pairs comprising a first polypeptide and a second polypeptide which method comprises:

(i) expressing a plurality of first polynucleotides to produce a plurality of first polypeptides and a plurality of second polynucleotides to produce a plurality of second polypeptides such that each first and second polynucleotide which directs expression of a corresponding first and second polypeptide are associated within the same unit displaying said first polypeptide

(ii) (ii) contacting said plurality of first polypeptides with a plurality of second polypeptides, wherein said second polypeptides are multivalent and wherein the contacting takes place under conditions that permit multivalent binding any of the second polypeptides to one or more of the first polypeptides.

(iii) selecting one or more interacting pairs of a first polypeptide that bind to a second polypeptide;

(iv) isolating the corresponding first and second polynucleotide(s) that encode(s) the one or more first and second polypeptides selected in step (iii).

Typically said second polypeptide is multivalent by virtue of being expressed as a fusion protein to a third polypeptide which multimerises.

As above, the second polypeptide may be present as one of a plurality of second polypeptides, which will typically be expressed from a plurality of second polynucleotides.

The methods of the invention are particularly suited to being implemented using phage display systems. Accordingly, the conditions that permit multivalent binding of the second polypeptide to one or more of the first polypeptides may comprise expressing the first polypeptides using phage display such that more than one molecule of a first polypeptide is present on the surface of the phage.

A consequence of the use of phage display is also that each first polynucleotide (including a f1 phage origin) which directs expression of a corresponding first polypeptide is associated with said first polypeptide since the polynucleotide is within the phage particle, and the polypeptide is displayed on the surface of the phage. Thus, once a first polypeptide of interest has been identified, the corresponding polynucleotide can easily be isolated.

Another consequence of the use of phage display is also that first and second polynucleotides (both including a f1 phage origin) which direct expression of the corresponding first and second polypeptides are associated with said first polypeptide since both polynucleotides are packaged within the phage particle, and the first polypeptide is displayed on the surface of the phage. Thus, once a first and second polypeptide combination of interest has been identified, the corresponding polynucleotides can easily be isolated.

The method of the invention can also be used to identify first and second interacting molecules whose interaction is modulated by the presence of absence of a third molecule. Accordingly, the above methods of the invention may comprise contacting said first molecule(s) and second molecule(s) in the presence or absence of a third molecule or plurality of third molecule, wherein binding of one or more first molecules and one or more second molecules is modulated by one or more third molecules.

The present invention also provides a first molecule, a second molecule and/or a third molecule identified by any of the methods of the invention.

Also provided is a kit comprising a plurality of first molecules and at least one second molecule for use in the methods of the invention.

Description of the Figures

Fig. 1. General scheme of Selection by avidity capture (SAC). (1) A receptor polypeptide fused to g3p and a ligand polypeptide (fused to a multimerizing domain, e.g. GST) are co-expressed in the same cell and exported to the periplasm, where (2), they associate to form a multivalent high avidity complex that is incorporated into nascent phage particles. (3) Phages bearing a cognate interaction complex are captured on a solid support via the GST domain. Because both plasmids contain a f1 packaging origin, both are packaged into phage particles. (4) Selected phage are plated for double antibiotic resistance and arrayed. Receptor and ligand proteins are coexpressed and cognate pairs detected (e.g. by capture of

the GST-ligand fusion protein with an anti-GST IgG and detection of cognate receptor binding using an anti-receptor-HRP conjugate).

Fig. 2. Effect of multivalency. The cognate receptor-ligand pair (scFv anti-BSA 13CG2 x GST-Protein L B1 domain) is rescued with two different helper phages. ELISA signal is plotted against phage titer. Phage are diluted stepwise by a factor of two, starting from 10^{12}cfu/ml (R408, filled circles) and 10^{11}cfu/ml (R408Δg3p, open squares). Phages rescued with R408 are predominantly monovalent, leading to inefficient avidity complex formation. Phages rescued with R408Δg3p are multivalent and avidity complexes are readily formed.

Fig. 3. Interaction matrix ELISA of avidity complexes on phage. Receptors and KD of cognate interactions: A: Fab 9E10 (anti-c-myc, KD= 80nM (Schiweck, *et al.* (1997) *FEBS Lett.* 414, 33-38)), B: FKBT- 12 (KD(FKBP-rapamycin, FRAP)= 2nM (Chen, *et al.* (1995) *Proc. Natl. Acad. Sci. USA* 92, 4947-4951)), C: FRAP (see B), D: scFvs M12 (anti-M, KD= 21nM (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994)), E: T14 (anti-T, KD= 5μM), F: D4 (anti-D (not determined)), G: c-Abl SH3, (we tested binding to two proline-rich target peptides, p41: KD(p41)= 1.5μM, and 3BP1: KD(3BP1)= 35μM (Pisabarro, *et al.* (1998) *J. Mol. Biol.* 281, 513-521)), H: scFvMab32 (anti- huTNFα, KD= 26 nM (Jespers, *et al.* (1994) *Bio/Technology* 12, 899-903)) are combined with ligands (1: GST-c-myc, 2: GSTFRAP, 3: GST-FKBP-12, 4: GST-M, 5: GST-T, 6: GST-D, 7: GST-41p and GST- 3BP1 peptide (shaded), 8: anti-Hen Egg Lysozyme (HEL) FvD1.3-huTNFα fusion protein (Holliger, P. (1994), PhD thesis, (ETH Zürich)) 9: anti-HEL diabody HyHEL10/5-c-myc (Holliger, *et al.* (1997) *Nat. Biotechnol.* 15, 632-63), 10: GST-Protein L B1 domain, KD(huVk1)= 130nM (Beckingham, *et al.* (1999) *Biochem J.* 340, 193-199)). SAC phage were rescued and all combinations assayed by ELISA (using either anti-GST IgG (rows 1-7, 10) or HEL (rows 8, 9) for capture of phage bearing avidity complexes).

Fig. 4. Stability of avidity complexes on phage. Dependence of avidity complex formation between FKBP-12 on phage and GST-FRAP on the extraneous mediator rapamycin and avidity complex stability is assayed by ELISA. Phage are rescued in the presence (+) or absence (-) of rapamycin, or are precipitated prior to the addition of rapamycin (PEG/+)

and captured with anti-GST IgG. Avidity complex formation is strictly dependent on rapamycin. (B) Once formed, FKBP-12 x GST-FRAP avidity complexes on phage are stable to repeated precipitation with polyethylene glycol (PEG).

Fig. 5. Results of SAC selection. Array screen of 3072 double spotted clones before and after one round of SAC selection. Clones were arrayed in a 4 x 4 pattern (see expanded panel). (A, B) Screening for cognate binding pairs before (A) and after (B) selection. Cognate interaction pairs are captured with anti-GST IgG and detected using Protein L-HRP. Selected clones are circled. Positive controls GST–M x anti-M scFv M12 and GST–T x anti-T scFv T4 are boxed. (C) Direct screen for antigen M specific binders after selection.

Detailed description of the invention

Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art (e.g., in cell culture, molecular genetics, nucleic acid chemistry, hybridisation techniques and biochemistry). Standard techniques are used for molecular, genetic and biochemical methods (see generally, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2d ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. and Ausubel *et al.*, Short Protocols in Molecular Biology (1999) 4th Ed, John Wiley & Sons, Inc. which are incorporated herein by reference) and chemical methods.

A. Molecules

The methods of the invention may be used to identify interactions between a wide variety of different types of molecules. Interactions, such as binding or association may be via covalent bonding, via ionic bonding, via hydrogen bonding, via Van-der-Waals bonding, or via any other type of reversible or irreversible association. However, typically, the interaction is of a reversible nature such as via ionic bonding, hydrogen bonding and/or Van-der-Waals bonding. Nonetheless, irreversible association may be desirable and can be

achieved by, for example, cross-linking by chemical or physical (e.g. UV cross-linking) means.

The methods of the invention may be also used to identify interactions resulting as part of or as a result of catalytic activity of one of the partners, including substrate or product binding in by a catalyst or covalent bond formation as a result of catalysis with a suicide inhibitor.

The term 'molecule' is used herein includes any atom, ion, molecule, macromolecule (for example polypeptide), or combination of such entities. Molecules used in the methods of the invention may be free in solution (including within any cellular compartment), or may be partially or fully immobilised. They may be present as discrete entities, or may be complexed with other molecules. Preferred molecules, particularly in the case of the first and/or second molecules, according to the invention include polypeptides.

Polypeptides include antibodies, antigens, enzymes, ligands such as growth factors, receptors such as cell surface receptors, DNA binding proteins such as DNA repair enzymes, polymerases, recombinases and transcription factors, and structural proteins. Polypeptides also include fragments of the above.

Polypeptides used in the methods of the invention may be provided as is, i.e. already synthesised by recombinant or chemical means. Alternatively, they may be provided by way of polynucleotides that express the polypeptides under suitable conditions. The polynucleotides may be present within a biological organism such as a bacterium, eukaryotic cell or virus. The nucleotides may also be provided as naked nucleic acids, typically as part of a vector such as a plasmid vector.

One particularly suitable system for providing polypeptides for use in the methods of the present invention are display methods, in particular phage display where the polypeptides are presented as integral parts of the envelope proteins on the outer surface of bacteriophage particles.

Other molecules which may be used in screens include nucleic acid molecules such as DNA or RNA, and libraries of peptide mimetics, defined chemical entities and natural product libraries.

The first and/or second molecules may comprise a tag or moiety to assist in purifying complexes formed between first and second molecules, and/or to assist in recovering first and/or second molecules for analysis and identification. Examples of tags include biotin, GST or a myc epitope tag.

Typically, the methods of the invention are used to screen a plurality of first molecules for interaction with a second molecule or plurality of second molecules. The second molecule is often termed a "bait' molecule. The plurality of first molecules whose interaction with the second molecule is being tested are often termed "prey" molecules. However, the second molecules may also be provided as a plurality of second molecules in a two-way screen of both "bait" and "prey" libraries.

The term "multivalent complex" as used herein means a molecular complex comprising (i) at least two molecules of a first molecule or at least one molecule of a multivalent first molecule and (ii) a multivalent second molecule (either multivalent by itself or rendered multivalent by fusion of an appropriate multimeric domain), wherein the at least two molecules of a first molecule are interacting with the second molecule or the at least one molecule of a multivalent first molecule is interacting with the second molecule via at least two valencies.

Thus, in the methods of the present invention, the second molecule is multivalent, which may, for example, be an inherent feature of the molecule (for example, haemoglobin is tetravalent with respect to oxygen binding) or as a result of suitable modifications to a monomeric molecule to allow the molecule to form multimers). The term multivalent includes divalent and higher orders of valency. The term "multivalent" means that the molecule has multiple sites for interaction with other molecules. By way of example, valency with respect to antibody molecules means the number of antigen binding sites on the molecule. IgM has 10 sites and thus has a valency of 10.

The second molecule is typically a multimer, such as a dimer, or capable of forming a homomultimer under the reaction conditions used in the methods of the invention. Alternatively, the second molecule may be a heteromultimer. It is preferred to use second polypeptides that have been engineered to be multimeric but whose constituent subunits do not normally form multimers. This may be achieved by chemical linkage of two or more molecules to form a covalently linked multimer. Alternatively, in the case of polypeptides, the molecules may be expressed as a fusion to a third polypeptide, the third polypeptide forming homomultimers which effectively results in multimerisation of the fused second polypeptide sequence. An example of a suitable third polypeptide (referred to herein as a "hook" polypeptide) is glutathione-S-transferase (GST), which forms dimers. Preferred hook polypeptides provide for good levels of expression of soluble products. In one embodiment, hook polypeptides that are suitable for secretion into the bacterial periplasm.

Where the second molecule is interacting with at least two distinct first molecules, the first molecules may be of the same or a different molecular species, preferably the same molecular species. Thus in the context of the present invention it is preferred that a multivalent molecule has two or more sites that are capable of binding simultaneously two or more other molecules or subunits that are identical.

Thus, by way of example, the second molecule may be a dimer with a binding domain on each monomeric subunit. The first molecule may be a molecule that can bind to the binding site on each monomeric subunit of the second molecule. The resulting complex therefore comprises one second molecule and two first molecules. It is not necessary for the two first molecules to interact with each other: the binding of both to the second molecule results in a more stable complex compared with one first molecule bound to the second molecule.

In a highly preferred embodiment, "multimers" in accordance with the invention does not include divalent IgG antibodies. Thus, a multimer is advantageously a molecule with a valency of two or more, other than IgG, any derivative thereof or other divalent immunoglobulin. Preferably, a multimer refers to a molecule with a valency of three or more.

In some embodiments, it may be desirable to engineer the first molecules such that they are multivalent, for example form multimers. Again, in the case of first polypeptides, the molecules may be expressed as a fusion to a fourth polypeptide (also referred to as a hook polypeptide), the fourth polypeptide forming homomultimers which effectively results in multimerisation of the fused first polypeptide sequence. The third polypeptide and fourth polypeptide should be selected so as avoid interactions solely between them in the absence of interactions between the first and second polypeptides otherwise false positives may result. The third and fourth polypeptides may be identical , e.g. GST.

Pluralities of first and second molecules include combinatorial libraries of chemical entities and randomised peptide libraries. Pluralities of first and second polynucleotides used to produce pluralities of first polypeptides and second polypeptides include cDNA expression libraries, exon trap libraries and randomised libraries. Methods for the production of libraries encoding randomised polypeptides are known in the art and may be applied in the present invention. Randomisation may be total, or partial; in the case of partial randomisation, the selected codons preferably encode options for amino acids, and not for stop codons. Libraries may also be created by DNA shuffling and molecular breeding methods (Stemmer 1994, Nature 370(6488): 389-91, Crameri 1998, Nature 391(6664): 288-91, Minshull 1999 Curr Opin Chem Biol 3(3): 284-90). Mutagenesis may be performed at the nucleic acid level, for example by site-directed or random mutagenesis of known gene sequences.

In a particularly preferred embodiment suited to high throughput screening, the first (and second) molecules are polypeptides encoded by first (and second) polynucleotides and the polynucleotides are associated with the corresponding polypeptides such that when a first (or a pair of first and second) polypeptide is selected during the screening methods of the present invention, the polynucleotide sequence that encodes the selected polypeptides is physically linked, or in close proximity, and can easily be recovered and, for example, sequenced to determine the identity of the selected polypeptides.

One way of achieving this is to provide the plurality of first (and second) polynucleotides (or a polynucleotide encoding both first (and second) polypeptides) in a compartment such that there is on average only one polynucleotide per compartment. The compartment

comprises the necessary reagents to allow for transcription/translation of the polynucleotide to produce the polypeptide. The use of phage display is a particularly suitable means of compartmentation since the nucleotide sequence is contained within the phage and the polypeptide is displayed on the surface of the phage where it can interact with a second molecule – see below.

However, other means that allow a first nucleotide to remain in association with a corresponding first polypeptide may also be used. Examples include in vivo and in vitro display technologies such as ribosome display (Mattheakis *et al.*, 1994 (Proc Natl Acad Sci U S A, 91, 9022-6) or puromycin-RNA linkage (Roberts & Szostak 1997, Proceedings of the National Academy of Sciences of the United States of America 94(23): 12297-302), combined with *in vitro* expression in suitable compartments, such as in water-in-oil emulsions (see WO 99/02671 and WO 00/40712), vesicles or simply enclosed in diffusion limiting materials such as gels.

The first and second polypeptides or plurality of polypeptides may be encoded on different nucleic acid molecules or a first polypeptide and a second polypeptide may be present as part of the same nucleic acid molecule.

The advantage of having the first polynucleotides and second polynucleotides as separate polynucleotide libraries is that combinatorial diversity can be produced by library "crossing" as well as providing the freedom to recombine different libraries without recloning. In phage display, combinatorial infection may be used to recombine bait and prey libraries. The assembled combinatorial diversity could then be "locked in" on a single replicon using, for example, a site-specific recombination process such as the cre-lox recombination system described previously (Waterhouse, 1993). However, the reassortment of two independent replicons is preferred because it should allow the continuous introduction of molecular diversity by reshuffling of libraries without the need for recloning or recombination.

However, where *in vitro* technologies such as ribosome display (Mattheakis *et al.*, 1994 (Proc Natl Acad Sci U S A, 91, 9022-6), Hanes, J. & Pluckthun, A. (1997) *Proc. Natl. Acad. Sci. USA* 94, 4937-4942) or puromycin-RNA linkage (Roberts, R. & Szostak, J.

(1997) *Proc Natl Acad Sci USA* 94, 12297-12302) are used, typically the first and second polynucleotides would be combined on the same template by ligation and co-expressed by *in vitro* translation in suitably non-communicating compartments, e.g. water-in-oil emulsion droplets (Tawfik & Griffiths 1998, Nature Biotechnology 16(7): 652-6).

In a particularly preferred embodiment of the present invention, the first polypeptides are produced using phage display, where the first polypeptide is expressed as a fusion protein with the a coat protein of bacteriophage, such as the minor coat protein pIII (g3p) of a filamentous bacteriophage e.g. bacteriophage M13, bacteriophage Fd etc. As a result it is displayed on the capsid of the bacteriophage is therefore accessible to the exterior environment. Detailed methodology for phage display is known in the art and set forth, for example, in US Patent 5,223,409; Choo and Klug, (1995) Current Opinions in Biotechnology 6:431-436; Smith, (1985) Science 228:1315-1317; and McCafferty *et al.*, (1990) Nature 348:552-554; all incorporated herein by reference. Vector systems and kits for phage display are available commercially, for example from Pharmacia.

When standard helper phages such as VCS M13, KO7, R408, which produce their own coat protein (in these cases g3p), are used for rescue, phage display on g3p is mostly monovalent, i.e. a majority of phage particles will display one or less fusion proteins. However, the valency of display can be increased by the use of helper phages lacking a coat protein gene (for example, Δg3p), such as those recently described (Rakonjac, 1997 Gene 198(1-2): 99-103). This ensures that all the g3p coat protein present on the surface of the phage derive from the g3p coat-protein-first polypeptide fusion protein produced by the phagemid. Notwithstanding proteolysis, this ensures that at least two or more first polypeptides are displayed on the tip of the phage particle, thus providing suitable conditions to allow for a multivalent interaction with a second molecule/polypeptide.

Thus, by way of example, a combination of Δg3p helper phage rescue, prey-g3p fusion polypeptides (first polypeptides) and a bait polypeptide (second polypeptide) fused to a multimersising "hook" such as GST may be employed. In this way, the prey-g3p fusion polypeptide is displayed in multiple copies on the phage surface and the bait polypeptide forms a dimer via the "hook". This allows divalent binding of two prey-g3p fusion polypeptides to the dimerised bait polypeptide. The combination of multivalent display and

dimeric hook allows formation of multivalent interaction and a long-lived interaction complex even for modest affinities. In the case where the bait polypeptide itself is multimeric, a "hook" polypeptide may not be required.

B. Screening methods

In its simplest format, an interaction between a first molecule and a second molecule is determined by incubating a candidate first molecule and a candidate second molecule such that an interaction may occur between them and then determining whether such an interaction has taken place. A variety of techniques may be used to determine whether an interaction has taken place.

Interaction of the first molecule with the second molecule may be assessed by any suitable means known to those skilled in the art. Interactions may be assessed with both the first and second molecule in solution. However, many techniques for detecting interactions rely on one of the components being immobilised to a solid phase.

Thus, in a number of methods, the first or second molecules or plurality of molecules are typically immobilised onto a solid phase. The substrate may be porous to allow immobilisation within the substrate or substantially non-porous, in which case the molecules are typically immobilised on the surface of the substrate. The solid substrate may be made of any material to which the molecules can bind, either directly or indirectly. Examples of suitable solid substrates include flat glass, silicon wafers, mica, ceramics and organic polymers such as sepharose and plastics, including polystyrene and polymethacrylate. It may also be possible to use semi-permeable membranes such as nitrocellulose or nylon membranes, which are widely available. The semi-permeable membranes may be mounted on a more robust solid surface such as glass. The surfaces may optionally be coated with a layer of metal, such as gold, platinum or other transition metal. A particular example of suitable solid substrate are the commercially available BiaCore™ chip (Pharmacia Biosensors), beads such as sepharose beads and plastic microtitre plates.

The solid substrate may be a material having a rigid or semi-rigid surface. In the case of "chips" at least one surface of the substrate will be substantially flat, although it may be desirable to provide physically separate regions for different molecules with, for example, raised regions or etched trenches. It is preferred that the solid substrate is suitable for the high density application of molecules in discrete areas of typically from 50 to 100 μm, giving a density of 10000 to 40000 cm^{-2}.

Flat solid substrates are conveniently divided up into sections. This may be achieved by the use of multiwell plates or in the case of chips, techniques such as photoetching, or by the application of hydrophobic inks, for example teflon-based inks (Cel-line, USA). Discrete positions, in which each different molecule is located may have any convenient shape, e.g., circular, rectangular, elliptical, wedge-shaped, etc.

Attachment of the molecules to the substrate may be by covalent or non-covalent means. The molecules may be attached to the substrate via a layer of molecules to which the first or second molecules bind. For example, the first or second molecules may be labelled with biotin and the substrate coated with avidin and/or streptavidin. A convenient feature of using biotinylated molecules is that the efficiency of coupling to the solid substrate can be determined easily. Other methods for attaching molecules to the surfaces of solid substrate by the use of coupling agents are known in the art, see for example WO98/49557.

Binding of first molecules to immobilised second molecules (or vice versa) may be determined by a variety of means such by the use of labelled molecules, such as epitope tagged polypeptides or polypeptides labelled with fluorophores such as green fluorescent protein. Binding of epitope tagged polypeptides is typically assessed by immunological detection techniques where the primary or secondary antibody comprises a detectable label. A preferred detectable label is one that emits light, such as a fluorophore, for example phycoerythrin.

Alternatively, the association could be monitored by fluorescent resonance energy transfer (FRET). In this case, the first molecule could be labelled with a donor fluor, and the second molecule could be labelled with a suitable acceptor fluor. Whilst the two entities are separated, no FRET would be observed, but if association (binding) took place, then

there would be a change in the amount of FRET observed, this allowing assessment of the degree of association.

Other detection techniques that do not require the use of labels include optical techniques such as optoacoustics, reflectometry, ellipsometry and surface plasmon resonance (SPR) – see WO97/49989.

In a preferred embodiment wherein the first molecules, for example, are phage displayed polypeptides, interactions between said molecules to form a multivalent complex may be assessed by eluting those phage which bind to an immobilised second molecule, and infecting logarithmic phase *E.coli* TG1 cells. The presence of infective particles eluted from the immobilised second molecule indicates that association of a first polypeptide with the second molecule has occurred. Alternative methods for detecting association between first and second molecules include phage ELISA.

Immobilisation of first or second molecules to a solid phase may take place after the first and second molecules are contacted such that multivalent complex formation occurs entirely in solution. Subsequent immobilisation has the effect of a purification/selection step to remove unbound molecules. For example, in the case of phage display, the bait molecule (second molecule) may be incubated with a library of phage and then the reaction mixture applied to an affinity column which comprises a ligand for the second molecule (or the hook polypeptide fused to the second molecule as appropriate). Phage that express prey polypeptides (first polypeptides) that do not bind to the bait molecule are washed through the column leaving only phage that express prey polypeptides that have formed a multivalent complex with the bait molecule immobilised to the affinity matrix. These phage may subsequently be eluted and the identity of interacting prey polypeptides identified, typically, by isolating individual phage clones and sequencing the corresponding first nucleotide encoding the first polypeptide.

In a preferred embodiment, a proteolytic cleavage site may be introduced between the hook polypeptide and the first or second polypeptide. Phage bound to the column by an interaction with an immobilised molecule may be released by cleavage of the phage displayed polypeptide. The hook polypeptide remains attached to the column and only the

bait protein (and bound phages) are eluted. This avoids the isolation of phage displaying prey molecules that interact with the column or the capture ligand or the hook domain bound to the capture ligand.

Generally, in the screening methods of the invention, at least the first molecules are provided as a plurality of first molecules. The use of phage display or in vitro "chip" technology allows large number of first molecules to be screened. Although the methods of the invention allow interactions to be detected which may normally be too weak to be detected using normal methods, it is likely that normally robust interactions will also be strengthened by virtue of the increased valency of the interaction. To simplify the analysis, it may be desirable to remove the higher affinity interactions in a pre-screening step. In the case of phage display, a screen could be carried out under conditions where multivalent interactions are reduced and affinity purification carried out to remove phage that express first polypeptides that bind to the second molecule. The flowthrough phage may then be used in a screen under conditions that promote multivalent interactions. Similar prescreens may be performed in non-phage systems.

Once a multivalent complex between a first molecule and a second molecule has been identified, typically the constituents of complex are then characterised. In this case of phage display, this can be achieved by isolating phage that express a component(s) of the complex and cloning and sequencing the nucleotide sequence(s) that encodes that component. Where *in vitro* methods are used where the nucleotide sequence that encodes a first and/or second polypeptide is associated with the polypeptide (e.g. ribosome display), then again, the components can be characterised by cloning and sequencing the relevant polynucleotide.

Where non-polypeptides or polypeptides that have not been produced recombinantly are screened, methods of identifying the components of the complex include mass spectroscopy and the use of molecular identification tags (for example in the case of combinatorial libraries of defined chemical entities).

It may be desirable to cross-link any multivalent complex using suitable chemical reagents or UV light for example, particularly where characterisation of the molecules is to be

performed using mass spectroscopy or where fleeting interactions need to be further stabilized.

In addition to two-way screens for interacting first and second molecules, the methods of the invention may be used to screen third molecules for the ability to modulate interactions between first and second molecules, for example the ability to enhance or reduce/inhibit an interaction between a first molecule and a second molecule. The methods employed are typically the same as described above. However, contacting of the first molecule and second molecule is performed in the presence or absence of a candidate third molecule or plurality of third molecules. Where an effect is seen on the interaction between the first and second molecule in the presence of the third molecule compared with the absence of the third molecule then the third molecule is selected and characterised. It may be desirable to screen pluralities of third molecules in pools and then deconvolute any given pool that shows an effect to identify the particular molecule within that pool that is having the effect. This enables larger numbers of compounds to be screened at a time.

C. Uses

The screening methods of the present invention may advantageously be used to identify interactions between molecules, particularly those among extracellular proteins, which frequently (especially those with disulfide bridges), do not attain full functionality when expressed in the reducing environment of the cytoplasm, where previously known techniques such as yeast two-hybrid screens and protein-complementation assay (PCA) approaches operate.

The screening methods of the present invention may be used for simultaneous interaction screening or selection e.g. for the isolation of specific antibodies and their antigens. Individual binding pairs could be identified and confirmed by de-convolution and array screening. The methods of the present invention may have a range of uses, including the parallel selection of many antibody-antigen pairs or the co-evolution of antibody-antigen interactions (e.g. for affinity maturation). The antibody-antigen interaction matrices have potential as tools for a better understanding of antibody-antigen interactions (e.g. "educated" libraries) as well as for expression profiling in research and diagnosis. Antibody

libraries from patients together with appropriate antigen libraires prepared from pathogens or diseased tissues (tumor biopsies, auto-immunity targets), may be selected to pinpoint immune response targets, antigenicity etc. for improved treatment or vaccine design. More generally, the methods of the present invention may be used to screen for ligands for orphan receptors or evolve agonistic ligands. It may also be possible to employ the methods of the present invention in generic library Vs library screens, identifying interacting pairs from molecular repertoires.

The invention may be applied in *in vivo* systems, including extracellular environments such as phage display, using phage such as M13, or phagemid, in accordance with conventional procedures. Intracellular *in vivo* environments are also accessible, using for example lambda phage display or T7 display techniques, or other intracellular expression systems.

The invention may be configured for *in vitro* use, for example using ribosome display or puromycin display techniques. Preferred is the use of compartmentalisation, for example in emulsions. Techniques for emulsification and selection are described herein.

Moreover, the invention may be exploited using screening systems other than genetic systems; for example, the use of mass spectrometry or surface plasmon resonance to detect binding is known in the art.

Because of the biases that multiple rounds of selection inevitably introduce, the methods of the present invention may be most powerfully employed in conjunction with array screening technologies (Holt et al 2000 Nucleic Acids Research 28(15): E72, De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994). A single round of selecting using the methods of the present invention should be sufficient to reduce diversity and enrich interacting pairs such that positives could be identified directly by array screening.

The phage display embodiment of the methods of the present invention is the only two-hybrid type system that operates extracellularly in the bacterial periplasm. Its periplasmic location should make it particularly suited for the analysis and selection of the interactions among extracellular proteins as well as the influence of small molecular compounds on defined receptor-ligand interactions or interaction matrices. Utility for drug-discovery

applications is enhanced by the periplasmic location as some compounds equilibrate poorly across the cell membrane into the cytoplasm, while molecules of up to 10kDa (Chen 2001 Nature Biotechnology 19(6): 537-42) can enter the periplasmic space. Being based on prokaryotic biology this embodiment may be particularly suited for the genomics analysis of the emerging microbial genomes, many of which are important pathogens.

Further advantages of using the phage-based prokaryotic embodiment of the methods of the invention is that repertoire sizes in yeast are substantially smaller than in *E. coli*. Thus, the numerical advantage of the larger repertoires achievable in *E. coli* might make phage-based methods of the present invention particularly suited for the genomics analysis of more complex genomes.

Detection of the avidity complex of the Abl SH3 domain with the 3BP-1 peptide (35μM) (Fig.3) suggests that interactions with affinities as low as 100μM should be selectable by SAC. Even lower affinity interactions may become accessible by the use of higher valency fusion tags, higher valency display (e.g. g8p (2800 copies per phage virion) or through the covalent stabilization of fleeting interactions, e.g. the formation of disulphide bridges or the addition of chemical crosslinking or bridging agents.

Conversely, capture of receptor-ligand complexes on phage may be used to select for high affinity interactions using monovalent display and monomeric ligands. Precedents for this approach exist. Antibody Fab fragments can be displayed and selected on phage. Fabs are formed by the non-covalent association of antibody heavy and light chains (K_D= <1pM). Thus, affinities in this range should allow monovalent capture to work, while the affinity of the Protein L B1-Vκ1 interaction (K_D= 130nM) appears too low (Fig. 2).

Because of its two-replicon format, pre-existing phage display libraries can be applied directly to SAC with no need for recloning or alterations to existing protocols. Many phage display libraries are constructed in colE1 phagemid vectors and thus are compatible and can be readily combined with our GST-ligand expression vector. This allows the application of SAC to the parallel screening and selection of antibody-antigen interactions (as described here), to the generation of very large combinatorial antibody libraries or as a generic screen for protein-protein interactions. The example of the rapamycin–mediated interaction of

FKBP-12 and FRAP (Fig. 2) shows that not all of the interaction partners need to be proteinaceous. SAC may be useful for the screening of compounds (like rapamycin) that mediate receptor-ligand interactions or complement mutations in a ligand-receptor interface.

SAC also offers the possibility of "capture display", e.g. the capture of soluble expressed multimeric antibody fragments or other proteins by a generic ligand displayed on phage. Indeed, a stable avidity complex can be assembled between a triabody and the Protein L B1 domain displayed on phage.

EXAMPLES

Examples & Methods

Methods

DNA manipulation and protein expression.

Glutathione-S-transferase (GST) (*S. japonicum*) is expressed from pGEX4T-2 (Pharmacia) and purified on glutathione sepharose (Pharmacia). Recombinant cDNA clones: (D (human chloride ion current inducer protein), RZPD clone MPMGp800E04369Q3), (M (unknown function), RZPD clone MPMGp800B12492Q3), and (T (ubiquitin), RZPD clone MPMGp800D17184Q3) derived from the human foetal brain cDNA library hEX1 (Bussow, *et al.* (1998) *Nucleic Acids Res.* 26, 5007-5008) and are expressed and purified as described (Lueking, *et al.* (1999) *Anal. Biochem.* 270, 103-111).

Receptor proteins are cloned in and displayed using phagemid vectors: pHEN-1 (Hoogenboom, *et al.* (1991) *Nucleic.Acids.Res.* 19, 4133-4137), pH (pHEN-1 without amber codon and c-myc tag) and pIT-2 (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994). pHENS, in which the Ampicillin resistance gene (*bla*) of pHEN-1 is replaced with the Spectinomycin resistance gene (*Sp*), is constructed as follows: the *Sp* gene from pSC101 is amplified with primers 1: (5'-TCA GCG CAC GCT GAC GTC GTG GAA ACG GAT GAA GGC ACG AAC-3') and 2: (5'-GCC GCC CGG GCA GTC GAC TTA TTA TTT GCC GAC TAC CTT GGT GAT CTC GCC-3'), cut with AatII and ligated with

pHEN-1 that is cut with AatII and DraI. Anti-huTNFα scFv Mab32 (Jespers, *et al.* (1994) *Bio/Technology* 12, 899-903), anti-M scFv M12, anti-T scFv T15, anti-D scFv D4 and antibovine serum albumin (BSA) scFv13CG2 have been described (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-99).

Cloning of receptor proteins: Fab 9E10 is amplified from pASK88-9E10 (Schiweck, *et al.*, A. (1997) *FEBS Lett.* 414, 33-38) using primers VHbaSfi and CkfoNot (Hoogenboom, *et al.*, (1991) *Nucleic.Acids.Res.* 19, 4133-4137), cut with SfiI/NotI and cloned into vectors pH (pHEN-1 without amber codon and c-myc tag) or fdSN (Choo, Y. & Klug, A. (1994) *Proc Natl Acad Sci USA* 91, 11168-11172) cut with SfiI and NotI.

Phage fdSN scFv Mab32 (α–huTNF) has been described (Jespers, *et al.* (1994) *Bio/Technology* 12, 899-903).

FKBP-12 is amplified from pGST-FKBP-12 (Main, et al. (1999) J. Mol. Biol. 291, 429-444) using primers S1: (5'- CGC GAC GGC TCG CGG CCC AGC CGG CCA TGG CCC AGG GTG TGC AGG TGG AAA CCA TCT CC-3') and S2: (5'-GGC CGA ATT CTT ATG CGG CCG CTT CCA GTT TTA GAA GCT CCA CAT C-3'), cut with SfiI/NotI and cloned into pH cut with SfiI and NotI. The FKBP-12- rapamycin binding domain of human FRAP is amplified from pGST-FRAP using primers S3: (5'- CGC GAC GGC TCG CGG CCC AGC CGG CCA TGG CCG AGC TGA TCC GTG TGG CCA TCC TC-3') and S4: (5'- GGC CGA ATT CTT ATG CGG CCG CCT GCT TTG AGA TAC GGC GGA ACA CAT G-3') cut with SfiI/NotI and cloned into pH cut with SfiI and NotI.

c-Abl SH3 (Pisabarro, et al. (1998) J. Mol. Biol. 281, 513-52) is amplified using primers S5: (5'-GCC ACG GCC ATG GCC AAT GAC CCC AAC CTT TTC GTT GCA-3') and S6: (5'-GGC CGC AGC GCT GCG GCC GCA CTG TTG ACT GGC GTG ATG TAG TT-3'), cut with NcoI/NotI and cloned into pHEN-1 (Hoogenboom, et al., (1991) Nucleic.Acids.Res. 19, 4133-41) cut with NcoI/NotI.

Peptostreptococcus magnus protein L domain D1 (Bjorck, L. (1988) J Immunol. 140, 1194-1197; Beckingham, et al. (1999) Biochem J. 340, 193-199) is amplified using primers S7: (5'- ACT GCG GCC CAG CCG GCC ATG GCC GAA GTA ACA ATC AAA GCT

AAC-3') and S8: (5'- GTG ATG ATG ATG TGC GGC CGC TCC AGC AAA TTT AAT ATT TAA-3') cut with NcoI/NotI and cloned into pIT-2 (De Wildt, et al. (2000) Nat. Biotechnol. 18, 989-9) cut with NcoI/ NotI.

Cloning of ligand proteins:

The ligand expression vector is constructed as follows: The lac promoter, *pel*B leader and polylinker of pUC119mycHis6 (Low, *et al.* (1996) *J.Mol.Biol.* 260, 359-368) are amplified using primers 3: (5'-CGG TGG CTG CCA TCG ATG GCA ACG CAA TTA ATG TGA GTT AGC TC-3') and 4: (no. 1211 New England Biolabs) cut with ClaI and ligated into pACYC184 (New England Biolabs) cut with Cla/BsaBI to give p184lacP. f1 origin of pUC119 is amplified with primers 5: (5'-GCT GCC GAC TCG GCT AGC GAA TGG CGA ATG GCG CCT GAT GCG G-3') and 6: (5'- GCC GGG TCG CTT TAA 6 AGT GTT GGC GGG TGT CGG GGC TGG C-3') cut with NheI / DraI and cloned into p184lacP cut with NheI / XmnI to give p184f1lacP. GST is amplified from pGEX4T-2 using primers 7 (5'-CGC CGG GAC TCG CGG CCC AGC CGG CCA TGG CCC AGT CCC CTA TAC TAG GTT ATT GG-3') and 8 (5'- CTC CGG CTG CGG CCG CAG CCT CGA GCG GGA ATC CAC GCG GAA CCA GAA CTT CCA GAT CCG ATT TTG GAG GAT G-3') cut SfiI /NotI and ligated into p184f1lacP cut with SfiI and NotI to give p184GST. For the human brain cDNA antigens, we used a variant of GST with 3 surface Cys mutated to Ser (GST3S) (Tudyka, T. & Skerra, A. (1997) *Protein Sci.* 6, 2180-2187) as fusion tag. GST3S is amplified with primers 7 and 9: (5'-CGG CTC CCC AGT CGA CCC GGG AAT TCC TGG GGA TCC ACG CGG AAC CAG ATC CGA TTT TGG AGG ATG GTC GCC ACC-3') and cut with SfiI and SalI, antigen genes M, T and E are excised from the hEX1 cDNA library with SalI/NotI and both fragments are coligated into p184f1lacP1 cut with SfiI and NotI to give vectors p184GST3S-M, T and E. Other ligand proteins are amplified using specific primers and cloned into p184f1lacP.

For experiments with phage fd receptor display vectors ligand proteins are expressed from pUC19SN. It is necessary to use a vector lacking a f1 origin because the presence of two f1 phage origins within the same cell (especially on a high copy number plasmid) causes toxic effects due to interference. Helper phage do not suffer from this problem because their f1 origin is defective.

pUC19SN is constructed as follows: the polylinker of pUC119mycHis6 (Low, et al. (1996) J.Mol.Biol. 260, 359-368) is excised by cutting Hind3/EcoRI and ligated into pUC19 (New England Biolabs) cut with Hind3/EcoRI. HyHEL10/5 diabody (Holliger, et al. (1997) Nat. Biotechnol. 15, 632-6) is excised Hind3/XhoI from pCantab5E (Pharmacia) and cloned into pUC19SN cut with Hind3/XhoI. FvD1.3-huTNFα (Holliger, P. (1994), PhD thesis (ETH Zürich)) is excised Hind3/NotI from pUC119mycHis6 and cloned into pUC19SN cut with Hind3/NotI.

For experiments with phagemid receptor display vectors, ligand proteins are expressed from p184f1lacP (see below). Cutting sites for NcoI and EcoRI in the *cat* gene of p184f1lacP are removed using Quickchange mutagenesis (Stratagene) and primers S9: (5'-CTT CGC CCC CGT TTT CAC TAT GGG CAA ATA TTA TAC GC- 3'), S10: (reverse complement of S9), S11: (5'-CTG ATG AAT GCT CAT CCG GAG TTC CGT ATG GCA ATG AAA G-3') and S12: (reverse complement of S11) to give vector p184NE. The ligand expression vector p184NE comprises promoter and polylinker of pUC119mycHis6 (Low, et al. (1996) J.Mol.Biol. 260, 359-368), therefore c-myc tag is appended to ligand proteins cloned NotI into p184NE, unless a stop codon precedes the NotI site.

HyHEL10/5 diabody (Holliger, et al. (1997) Nat. Biotechnol. 15, 632-6) is excised Hind3/XhoI from pCantab5E (Pharmacia) and cloned into p184lacP cut with Hind3/XhoI.

FvD1.3-huTNFα (Holliger, P. (1994), PhD thesis (ETH Zürich)) is excised Hind3/NotI from pUC119mycHis6 and cloned into p184lacP cut with Hind3/NotI.

Anti-BSA triabody 13CG2 is generated from the anti-BSA scFv 13CG2 (De Wildt, et al. (2000) Nat. Biotechnol. 18, 989-94) by cutting out of the interdomain linker with XhoI/SalI and religation (Tomlinson, I. & Holliger, P. (2000) Meth. Enzymol. 326, 461-79). The triabody is excised Hind3/NotI from pIT-2 and cloned into p184lacP cut with Hind3/NotI. Protein L domain D1 (Bjorck, L. (1988) J Immunol. 140, 1194-1197; Beckingham, et al. (1999) Biochem J. 340, 193-199) is amplified using primers S13: (5'- GGG GCT GAG TCA CTC GAG GAA GTA ACA ATC AAA GCT AAC CTA-3') and S8: (5'- GTG ATG ATG ATG TGC GGC CGC TCC AGC AAA TTT AAT ATT TAA-3') cut with XhoI/NotI.

GST3S (a GST variant with 3 surface Cys mutated to Ser) is amplified with primers 7 and 9: (5'-CGG CTC CCC AGT CGA CCC GGG AAT TCC TGG GGA TCC ACG CGG AAC CAG ATC CGA TTT TGG AGG ATG GTC GCC ACC-3') and cut with SfiI and SalI and both fragments are co-ligated into p184f1lacP1 cut with SfiI and NotI. GST3S has been reported to be a superior fusion tag for periplasmic expression (Tudyka, T. & Skerra, A. (1997) Protein Science 6, 2180-7). In our hands, both variants (GST and GST3S) performed equally well in SAC. GST-FKBP-12 is amplified from pGST-FKBP-12 (Main, et al. (1999) J. Mol. Biol. 291, 429-444) using primers 11 and S2, cut with SfiI/EcoRI and cloned into p184NE cut with SfiI/EcoRI.

GST-FRAP is amplified from pGST-FRAP 12 using primers 11 and S4, cut with SfiI/EcoRI and cloned into p184NE cut with SfiI/EcoRI. Abl SH3 binding proline-rich peptides 41p (high affinity) and low affinity (3BP1) (Pisabarro, et al. (1998) J. Mol. Biol. 281, 513-521) are appended to GST as follows: GST is amplified from pGEX-4T2 (Pharmacia) using primers 11 and S14: (5'- GGC GAA TTC TTA AGG TGG GGG CGG AGG AGA GTA AGA GGG AGC GGA TCC ACG CGG AAC CAG-3') encoding 41p and with primers 11 and S15: (5'- GGC GAA TTC TTA AGG GGG AAG CGG AGG GGG CAT AGT GGG AGC GGA TCC ACG CGG AAC CAG-3') encoding 3BP1.

SAC phage growth:

Plasmids are co-transformed (or co-infected) into E.coli TG1 and plated on TYE/Ampicillin (Amp) (0.1mg/ml) / Chloramphenicol (Chl) (10μg/ml) / 5% Glucose plates (TYAC5 plates). Single colonies are picked and inoculated into 2xTY / Amp (0.1mg/ml) / Chl (10μg/ml) / 5% Glucose (TYAC5 medium) and grown overnight (ON). ON cultures are diluted 1/100 into fresh TYAC5 medium, grown for 1h at 37°C and R408Δg3p phage (Rakonjac, J., *et al.* (1997) *Gene* 198, 99-103) added to a final concentration of 10^9 plaque forming units (pfu) / ml. The culture is incubated standing at 37°C for 1.5h, spun down and resuspended in an equal volume of Terrific Broth / Carbenicillin (Sigma) (0.1mg/ml) / Chl (10μg/ml)/ 0.4% Glucose (including 1nM Rapamycin (Calbiochem) for FKBP-12 x FRAP) and grown ON at 25°C. Phage are isolated by polyethylene glycol (PEG) precipitation, titered and assayed by phage ELISA using standard protocols (Harrison, *et al.* (1996) *Meth. Enzymol.* 267, 83-109).

fdSN Fab910 and fdSN scFv Mab32 phage are similarly used to infect TG1 cells harbouring pUC19SN: HyHEL10myc or pUC19SN: FvD1.3-TNFα plasmids and plated on TYE / Ampicillin (Amp) (0.1mg/ml) / Tetracycline (Tet) (15μg/ml) / 5% Glucose plates. Single colonies are picked and inoculated into Terrific Broth / Amp (0.1mg/ml) / Tet (0.2μg/ml) / 0.4% Glucose and grown overnight at 30°C. Phage are isolated by polyethylene glycol (PEG) precipitation, titered and assayed by phage ELISA using standard protocols (Harrison, et al. (1996) Meth. Enzymol. 267, 83-109).

Array screening : The single framework (V3-23/DP-47, JH4b x O12/O2/DPK9, Jκ1) libraries I and J have been described (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994). Phages are produced from library I and J using standard protocols (Harrison, *et al.* (1996) *Meth. Enzymol.* 267, 83-109) and used to infect TG1 cells harbouring ligand expression vector p184GST-M or an equal mix of p184GST-D, M and T. Infected cells are plated on TYAC5 plates and grown ON at 30°C. Plates are scraped as described (Harrison, *et al.* (1996) *Meth. Enzymol.* 267, 83-109) and bacteria added to prewarmed TYAC5 medium to a final OD600 of 0.4. Phage particles are produced using R408Δg3p helper phage as described above. Phages are selected on glutathione sepharose (GS). GS is incubated for 2h in PBS, 2% Tween 20 at RT and washed 2x with PBS, 0.1% Tween20 (PBST). Phage are incubated for 15' in PBST prior to addition to GS, mixed for 1h with GS on a rotating platform, then washed (4x 25ml PBST). Phages are eluted by incubation of the resin with 1 mg/ml trypsin in PBS for 10 min. Eluted phages are used to infect TG1 and plated onto TYAC5 plates. Colonies are robotically picked and gridded.

Picking and gridding is performed as described (De Wildt, et al. (2000) Nat. Biotechnol. 18, 989-94). Briefly, selected clones are plated onto a large square agar plate (230 x 230 mm, Nunc plates containing TYE, 10 μg/ml chloramphenicol, 100 μg/ml ampicillin, 1% glucose). Single colonies are picked using a robotic colony picker/gridder sytem (Kbiosystems, Basildon, UK) into 384- well plates (containing TYE, 10 μg/ml chloramphenicol, 100 μg/ml ampicillin, 1% glucose, 8% glycerol, 75 μl/well) and grown overnight at 37°C. These are then gridded in a 4 x 4 pattern (such that each clone is gridded twice in each 4x4 grouping, see Fig.5) onto a large square agar plate covered with a nitrocellulose filter (Protran BA85; Schleicher & Schuell), that had been blocked in 2%

dried skimmed milk (Marvel) in PBS (MPBS) for 60 min at room temperature (RT), briefly washed in PBS and soaked in 2xTY. This blocking procedure prevents binding of recombinant proteins to this filter. The gridded plates are grown overnight at 37°C. In the meantime, a second nitrocellulose filter is coated overnight at 4°C in 80 ml of PBS with either 2.5 μg/ml of goat anti-GST antibody (Amersham Pharmacia Biotech), 100 μg/ml of bovine serum albumin (BSA) control antigen, or 50 μg/ml of purified recombinant D, M, T or recombinant GST control. We used goat anti-GST antibody as capturing reagent since protein L (used for detection) does not bind goat immunoglobulins. This second filter is blocked in 2% MPBS for 1 hr RT, washed 3x in PBS, soaked in 2xTY and transferred onto a large square plate (containing TYE, 100 μg/ml ampicillin, 1 mM isopropyl β-D-thiogalactoside (IPTG)). The first filter containing the grown colonies is transferred onto the plate covered with the second filter. These plates are incubated for 5 hr at 30°C to induce expression of GST-fusion and scFvs proteins. The top filter is discarded and the bottom filter washed two times with PBS/0.05% Tween (PBST) and then blocked with 2% MPBS for 30 min at RT, followed by two washes with PBST. To detect binding of scFvs, the filter is incubated with Protein L-HRP conjugate (Affitech, 1/2000) in 2% MPBS for one hour at RT. The filters are washed three times with PBST and developed using electrochemiluminescence (ECL) reagent. All incubations are performed in 50 ml of buffer on a gently agitating shaker.

Antibody purification and affinity determination. Selected scFv fragments are expressed and purified on protein-A sepharose (Pharmacia) as described (Tomlinson, I. & Holliger, P. (2000) *Methods Enzymol.* 326, 461-79) and binding affinities of scFv fragments are determined using plasmon surface resonance with BIAcore as described.

Example 1) Specific avidity complexes on phage fd.

Phage vectors (e.g. fdDOG) offer the advantages of polyvalent display. Thus as a first test of the technology a multivalent antibody-antigen complex is assembled on the tip of phage fd.

Two model interactions are tested between prey (antibody) and bait (antigen): the αTNF antibody Mab32 (displayed as a scFv fragment) with TNFα (trimeric, fused to a αFITC scFv or anti- Hen Egg Lysozyme (HEL) Fv D1.3 for capture on respective antigen) and the αcMyc antibody 9E10 (displayed as a Fab fragment) with the cMyc peptide fused to a Glutathione-S-transferase (dimeric; from *S. Japonicum*) hook domain. The bait proteins are cloned into pUC19 derived expression vectors and coexpressed in cells infected with the phage displaying the prey proteins (i.e. the antibodies).

1) In both cases specific interactions are reconstituted on the surface of the phage as judged by specific binding of the phage in ELISA to an interaction surface specified by the hook. Specifically Mab32 phage derived from cells expressing the anti-FITC scFv-TNF fusion protein bind to FITC-BSA and phage derived from cells expressing the anti-HEL Fv-TNF fusion protein bind to HEL and but not to an αGST antibody or BSA, while 9E10 phage derived from cells expressing the GST-myc bind to an αGST antibody but neither to FITC-BSA nor to BSA.

2) Phages derived from non-cognate pairings (e.g. Mab32 x GSTmyc) do not bind any of the capture antigens.

3) Antibody-antigen complexes are stable to repeated PEG precipitations suggesting that they are preassembled and stably associated with the phage particle and do not assemble on the ELISA plate surface from soluble bait and prey phage. Once assembled interaction-complexes can be stored at 4°C for weeks without loss of signal. This is significant, in particular for antibody-peptide interactions, as exposed peptide tags are particularly vulnerable to cleavage by residual proteases. In general, peptide tags are completely cleaved after a few days incubation in a crude protein preparation (e.g. culture supernatants, periplasmic preparations) at 4°C. The fact that antibody-peptide interactions remain stable for weeks suggests that the peptide is protected from proteolytic attack, presumably by being bound in the binding site of the antibody.

4) Filtration through a 0.45μM filter does not reduce specific ELISA signals, indicating that signals derive from particles not significantly larger than phage particles.

Example 2) Avidity complexes assembled using a phagemid system

Phagemid vectors are normally used for monovalent display as the commonly used helper phages like KO7 or R408 produce g3p in trans. Use of the recently described Δg3p helper phages R408 Δg3p and ΔK07 (Rakonjac, *et al.* (1997) *Gene* 198, 99-103) allows polyvalent display on phagemid vectors. In order to be able to use previously constructed phage display libraries (commonly cloned in colE vectors) as prey libraries we constructed a p15 origin based expression vector for prey proteins derived from pACYC184 (p184f1lacP and derivatives). In some experiments, a variant GST in which 3 surface Cys are mutated to Ser (GST3S) protein for better periplasmic expression is used as the hook protein.

Table 1 shows a list of model interactions demonstrated with this system. These include protein-protein as well as peptide-protein interactions as well as a model 3-hybrid interaction, i.e. the interaction between FRAP and FKBP-12 mediated by rapamycin (see below).

1) All the interactions listed in Table 1 are found to be specific, i.e. binding to solid support is strictly dependent on both the identity of the hook portion (e.g. GST) as well as the presence of cognate interaction partners or in the case of FKBP-12 and FRAP on the presence of the interaction mediator rapamycin. Phages derived from non-cognate pairings do not bind (see also Fig. 3).

2) As for the phage system (see example 1), phages derived from non-cognate pairings do not bind any of the capture antigens. Interaction complexes are stable to repeated PEG precipitations suggesting that they are preassembled and stably associated with the phage particle and did not assemble on the ELISA plate surface from soluble bait and prey phage.

3) As expected, stable interaction complexes can not be assembled efficiently using KO7 or R408 helper phages (which predominately produce monovalent phage particles) but require the Δg3p phages R408Δg3p and ΔK07 (see Fig. 2). Of these, R408Δg3p is the

preferred helper phage, because the p15 origins of KO7 and the pACYC184-based bait plasmid interfere with each other and lead to reduced phage titers.

4) Stability of such complexes depends on the valency of the interaction partners. Monovalent bait proteins do not give rise to stable interaction complexes even if the prey is displayed polyvalently on phage, while dimeric diabodies (displaying 2 cMyc peptide tags per molecule), together with Fab9E10 displayed on phage, lead to the assembly of stable interaction complexes (as judged by ELISA and PEG precipitation), monomeric Fv fragments (displaying 2 cMyc peptide tags per molecule) do not.

5) As for the phage system (see example 1), filtration through a 0.45μM filter do not reduce specific ELISA signals, indicating that signals derive from particles not significantly larger than phage particles.

Table 1: Model interactions demonstrated with the SAC system

bait (fused to dimeric GST hook, capture with anti-GST IgG))	prey (N-terminal g3p display)
CMyc peptide	Fab 9E10
scFv αFITC-cmyc (capture also with FITC-BSA)	Fab 9E10
Protein-L (B1 domain)	Several scFv fragments (see Fig. 3)
scFv αBSA 13CG2	Protein-L (B1 domain)
c-abl SH3 domain	3BP1 target peptide (35μM) 41P peptide (1.5μM)
antigen M (hu cDNA library)	αM scFv M12 (& selected scFvs)
Ubiquitin (antigen T)	αUbiquitin scFv T14 (& selected scFvs)
antigen D	αD scFv D4 (& selected scFvs)
FKBP-12 (+ rapamycin)	FRAP
FRAP (+ rapamycin)	FKBP-12

bait (multimeric)	prey (N-terminal g3p display)
Diabody D1.3/5-cmyc (dimer)	Fab 9E10
Diabody NQ11/5-cmyc (dimer)	Fab 9E10
Diabody HyHEL10/5-cmyc (dimer)	Fab 9E10
ScFv αFITC -TNFα (trimer)	scFv Mab32 (αTNF)
Fv D1.3-TNFα (trimer)	scFv Mab32 (αTNF)
triabody αBSA 13CG2	Protein-L (B1 domain)

Example 3 : A 3-hybrid interaction: a two-protein interaction mediated by a small molecule: FRAP-rapamycin-FKBP-12

The interaction between the FRAP domain of the Tol protein kinase (Brown 1994 Nature 369(6483): 756-8) and the FKBP-12 Pro-isomerase is mediated by the small cyclic peptide rapamycin. This interaction inhibits T-cell activation and rapamycin is an important drug for the prevention and treatment of graft vs. host disease in organ transplantation.

1) We found that as in its natural environment the interaction between FRAP and FKBP-12 (either GST-FRAP/FKBP-12-phage or GST-FKBP-12/FRAP-phage, see Table 1) is specific and strictly dependent on the presence of rapamycin (Fig. 4).

2) As for the other examples of interaction complexes (see Table 1), phages derived from non-cognate pairings or produced in the absence of rapamycin did not bind any of the capture antigens. Interaction complexes are stable to repeated PEG precipitations suggesting that they are preassembled and stably associated with the phage particle and do not assemble on the ELISA plate surface from soluble bait and prey phage.

3) Stable interaction complexes can not be assembled efficiently using KO7 or R408 helper phages but require the Δg3p phage R408Δg3p (Fig. 2).

4) If phages are produced and PEG precipitated in the absence of rapamycin, addition of even a large excess of rapamycin only results in the recovery of 1-5% of the ELISA signal of the cognate interaction (Fig. 4). This indicates (as would be expected) that the soluble interaction partner (either GST-FRAP or GST-FKBP-12), if not bound to the phage particle, largely remains in the supernatant and is lost during the PEG precipitation.

5) Filtration through a 0.45μM filter does not reduce specific ELISA signals, indicating that signals derive from particles not significantly larger than phage particles.

6) The interaction between GST-FRAP and FKBP-12-phage displays an IC_{50} value for rapamycin that is very close to the value measured for in-solution interaction (not shown).

Conclusion:

Taken together examples 1-3 establish:

I) interaction complexes can be assembled on the tip of filamentous phage particles

II) stability of such complexes depends on the valency of the interaction partners, i.e. avidity. Multivalency can be encoded in the nature of the hook protein (e.g. dimeric GST) or in the nature of the bait (TNFα, trimeric by itself). Prey proteins are displayed on phage, preferably in multiple copies per particle, preferably on g3p by the use Δg3p helper phages like R408Δg3p. Alternatively, they may be displayed on g8p or any other phage protein.

III) once assembled interaction complexes remain stable to high salt, PEG precipitation and prolonged storage at 4°C.

IV) interaction complexes are specific and can be detected by ELISA

V) complexes involving more than two interaction partners (e.g. 3-hybrid) can be assembled

Example 4: Detecting cognate protein-protein interactions.

The ability of SAC to detect a diverse range of receptor-ligand interactions including antibody-antigen interactions is investigated using a matrix screen comprising all possible permutations of ten ligand and eight receptor proteins (Fig. 3). The permutations tested are as follows, with reference to Fig. 3, receptors (prey proteins) columns A-H: A: Fab 9E10 (anti-c-myc, KD= 80nM (Schiweck, *et al.* (1997) *FEBS Lett.* 414, 33-38)), B: FKBT- 12 (KD(FKBP-rapamycin, FRAP)= 2nM (Chen, *et al.* (1995) *Proc. Natl. Acad. Sci. USA* 92, 4947-4951)), C: FRAP (see B), D: scFvs M12 (anti- M, KD= 21nM (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994)), E: T14 (anti-T, KD= 5μM), F: D4 (anti-D (not determined)), G: c-Abl SH3, (we tested binding to two proline-rich target peptides, p41: KD(p41)= 1.5μM, and 3BP1: KD(3BP1)= 35μM (Pisabarro, *et al.* (1998) *J. Mol. Biol.* 281, 513-521)), H: scFvMab32 (anti- huTNFα, KD= 26 nM (Jespers, *et al.* (1994) *Bio/Technology* 12, 899-903)) are combined with ligands (bait proteins), rows 1-10: (1: GST-c-myc, 2: GSTFRAP, 3: GST-FKBP-12, 4: GST-M, 5: GST-T, 6: GST-D, 7: GST-41p and GST- 3BP1 peptide (shaded), 8: anti-Hen Egg Lysozyme (HEL) FvD1.3-huTNFα fusion protein (Holliger, P. (1994), PhD thesis, (ETH Zürich)) 9: anti-HEL diabody HyHEL10/5-c-myc (Holliger, *et al.* (1997) *Nat. Biotechnol.* 15, 632-63), 10: GST-Protein L B1 domain, KD(huVk1)= 130nM (Beckingham, *et al.* (1999) *Biochem J.* 340, 193-199)). SAC phage are rescued and all combinations assayed by ELISA (using either anti-GST IgG (rows 1-7, 10) or HEL (rows 8, 9) for capture of phage bearing avidity complexes).

SAC detects interactions of a wide range of affinities (down to 35μM). Only the cognate interactions give rise to an ELISA signal, with exception of the anti-D scFv D4 (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994), which shows crossreactivity with huTNFα. The identification of unexpected cross-reactivities is important, in particular for application

of antibodies in therapy. The above example shows the potential of SAC matrix screening for detecting such cross-reactivities.

The high avidity interaction complexes assembled on the phage tip appear to be exceptionally stable. They survive multiple, sequential precipitations of the phage particles with polyethylene glycol (PEG) (Fig. 4) and are stable to storage at 4°C for several weeks. Phage bearing avidity complexes can also be filtered through 0.45μM filters without loss of either phage titer or ELISA signal, suggesting that avidity complexes do neither significantly increase phage size nor involve multiple phages. Furthermore, unlike in the SIP/SAP approach, phage infectivity and hence library size are not affected by the assembly of multimeric interaction complexes on the phage tip. PEG precipitation is also an efficient way to remove soluble ligand protein present in the culture supernatant from the phage preparation, as is demonstrated for the rapamycin-dependent interaction between FKBP-12 and FRAP (Chen, *et al.* (1995) *Proc. Natl. Acad. Sci. USA* 92, 4947-4951). If rapamycin is added during phage rescue or prior to phage precipitation with PEG, it drives association of FKBP-12 and FRAP. In the absence of rapamycin or if it is added after PEG precipitation, no (or much reduced) complex formation is observed (Fig. 4).

Example 5: Selection of cognate interactions by SAC.

The utility of SAC for the selection of cognate interactions is tested using an antibody-antigen interaction pair (human foetal brain antigen M fused to GST (GST-M) and its cognate anti-M antibody M12 (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994)) as a model. To aid selections a protease cleavage site is introduced between the GST and the ligand domain in the ligand expression vector p184GST to allow specific elution by protease cleavage. In order to distinguish between cognate and non-cognate phage, we constructed the Spectinomycin resistance (SpR) conferring phagemid: pHENS for display of scFv M12. As non-cognate phage we used the anti-BSA scFv 13CG2 cloned into the Ampicillin resistance (AmpR) conferring phagemid pHEN-1. Both phagemids are used to infect cells harbouring p184GSTM (which confers Chloramphenicol resistance (ChlR)). From these, phage are rescued using R408Δg3p (Rakonjac, J., *et al.* (1997) *Gene* 198, 99-103).

To test the ability of SAC to select for cognate interactions cognate phage (pHENS:scFv M12 x p184GST-M) (SpR/ChlR) is spiked into a 10^3-10^4-fold excess of non-cognate phage (pHEN-1:scFv 13CG2 x p184GST-M) (AmpR/ChlR). Using either capture by an anti-GST antibody (coated on plastic) or glutathione sepharose beads and specific elution of captured phage by protease cleavage, it is found that up to a 1000-fold enrichment can be achieved in a single round of SAC (as judged by comparing the ratio of SpR/ChlR to AmpR/ChlR phage before and after selection). Enrichment factors for SAC are thus comparable to those observed for standard phage display (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994).

Example 6: Selection of antibody-antigen pairs using SAC.

SAC is also applied to the isolation of antibody-antigen interaction pairs directly from a naive antibody scFv library. Because of its two-replicon format, SAC can be applied to any pre-existing display library constructed in standard phagemid vectors without the need for recloning.

Two synthetic, single human framework scFv libraries I and J are used (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994). Library I is first combined with GST-M and a single round of SAC performed, capturing cognate phage on glutathione-sepharose. Input and eluted phage are plated on Amp/Chl plates and 3072 colonies each are robotically picked and gridded for array screening. Cognate interactions are identified by capture of the SAC complex with an anti-GST antibody and detection of the antibody portion by Protein L. Specific binding of the antibodies directly to recombinant antigen M or non-specific binding to GST or an irrelevant antigen (BSA) is also investigated. Circa 60 positive clones are detected (Fig 5B) on the screen for the cognate interactions. A larger number (ca. 110, Fig 5C) appear to bind specifically to antigen M but are not (or only weakly) detectable on the cognate screen, suggesting that the cognate screen is more stringent. Coexpression of GST-M, followed by capture on anti-GST IgG may result in a lower local antigen density than direct coating. Nevertheless, antigen binding appears highly specific with no binders detected on the GST and BSA screens. Neither are any positive clones detected from screening the unselected library (Fig 5A) indicating an enrichment of

specific binders of at least a 100-fold. Comparable enrichment factors are obtained by conventional phage display.

30 clones are chosen (with a range of positive signals) from the cognate screen (Fig. 5B) and their specificity tested by ELISA against antigen M and control antigens (BSA, GST). A majority (21) is specific for M, while for the other 9 clones, no binding can be detected on any of the antigens tested. Because of the higher density of antigen coating and less stringent washing, array screening detects lower affinity interactions than ELISA as observed previously (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994). Sequence analysis of the 21 specific clones shows that they are all different but contain Serine-rich VH CDR3 sequences very similar to anti-M scFvs isolated by conventional phage display. Indeed, scFv M1/14 (Table 2) displays an identical VH CDR3 sequence (SSYS; but different VH CDR1 and VL sequences) as scFv M4 isolated previously by conventional phage display (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994).

Table 2. Antibody sequences of clones isolated using SAC and array screening

Clone	Library*	Antigen	CDRH2	CDRH3	CDRL2	CDRL3
Single antigen M						
M1/14†	DVT(I)	M	AITYSGAGTTYADSVKG	SSYSFDY	AASSLQS	QQDSYAPST
M1/15	DVT(I)	M	AITSSGADTTYADSVKG	AGYSFDY	CASSLQS	QQYAYCPGT
M1/16	DVT(I)	M	GITYTGTYTTYADSVKG	SSTGFDY	SASTLQS	QQSNATPGT
M1/17	DVT(I)	M	AISYTGNNTYYADSVKG	AGSSFDY	TASSLQS	QQDSANPAT
M1/18	DVT(I)	M	SIYSAGAGTYYADSVKG	AYGSFDY	SASTLQS	QQTNSGPAT
M1/19	DVT(I)	M	AISYTGNSTTYADSVKG	SYSGFDY	TASALQS	QQSDTNPGT
M1/20	DVT(I)	M	NITYSGTNTSYADSVKG	SYYAFDY	SASALQS	QQNDTTPYT
M1/21	DVT(I)	M	GITYAGAATTYADSVKG	ATYCFDY	YASALQS	QQTSYYPGT
M1/22	DVT(I)	M	YISYNGSYTNYADSVKG	SATSFDY	NASSLQS	QQNDSTPST
Mixture of antigens D, M and T						
M2/1	DVT(I)	M	TISYAGYATTYADSVKG	AASDFDY	NASSLQS	QQYASSPNT
M2/2	DVT(I)	M	SISYSGNTTSYADSVKG	VYCWFDY	NASSLQS	QQSNSYPGT
M2/3	DVT(I)	M	NISYSGSSTYYADSVKG	ASTSFDY	TASGLQS	QQDSDAPGT
M2/4	DVT(I)	M	GISYTGTSTYYADSVKG	SNSNFDY	SASKLQS	QQDNTDPST
M2/5	DVT(I)	M	GISYTGNSTTYADSVKG	SSYNFDY	AASNLQS	QQTAYDPYT
M2/6	DVT(I)	M	TISYNGGSTYYADSVKG	STSNFDY	AASTLQS	QQSTYNPDT
M2/7	DVT(I)	M	NITYTGTYTTYADSVKG	ASAGFDY	SASALQS	QQCSNTPCT
M2/8	DVT(I)	M	NISSYGTTTYYADSVKG	ANTYFDY	SASALQS	QQYSSNPNT
M2/10	DVT(I)	M	SISYTGSSTSYADSVKG	STYSFDY	SASSLQS	QQDTTSPCT
M2/13	DVT(I)	M	GISTDGTSTNYADSVKG	SSSAFDY	SASGLQS	QQAYSNPGT
M2/15	DVT(I)	M	GITYSGAYTSYADSVKG	SYSAFDY	AASALQS	QQSTGYPST
M2/18	DVT(I)	M	TISYSGNNTAYADSVKG	TGTYFDY	GASYLQS	QQYGDAPAT
M2/19	DVT(I)	M	GITYTGSSTSYADSVKG	YGAYFDY	NASGLQS	QQDSSTPST
M2/20	DVT(I)	M	NISYSGNSTGYADSVKG	ANYSFDY	SASDLQS	QQDSANPNT
M3/8	DVT(I)	M	NITYNGAYTTYADSVKG	AYTSFDY	AASSLQS	QQSYSTPNT
M3/14	DVT(I)	M	GITYAGNYTYYADSVKG	ASSSFDY	AASSLQS	QQDGANPST
M3/16	DVT(I)	M	AITYSGTSTTYADSVKG	STYSFDY	DASNLQS	QQDSNNPYT
M3/17	DVT(I)	M	SISYAGANTYYADSVKG	YSANFDY	AASSLQS	QQAGSSPGT
M3/21	NNK(J)	M	RIDRTGTHTLYADSVKG	TLTPFDY	MASSLQS	QQIPLSPQT

D2/4 DVT(I)	D	AITSAGSYTAYADSVKG	ANTSFDY	SASNLQS	QQTYAAPGT
D2/6 NNK(J)	D	SISSRGRSTGYADSVKG	TTHLFDY	AASSLQS	QQSYSTPNT
D2/7 NNK(J)	D	TIAGPGWRTHYADSVKG	SYPGFDY	GASLLQS	QQKNARPST
D2/8 NNK(J)	D	SIHNGGNRTSYADSVKG	PRVSFDY	HASTLQS	QQSERPPDT
D2/9 NNK(J)	D	TIGRTGSHTGYADSVKG	SLLGFDY	NASRLQS	QQRETFPRT
D2/10 NNK(J)	D	TIA--GKRTAYADSVKG	LTLNFDY	KASGLQS	QQANPLPHT
D2/11 NNK(J)	D	LIARTGRTTQYADSVKG	LIPTFDY	AASSLQS	QQHSGTPPT
D2/12 NNK(J)	D	AIQASGGRTIYADSVKG	AGNPFDY	EASSLQS	QQRRQTPYT
T2/12 DVT(I)	T	YIDSNGTNTAYADSVKG	TTDAFDY	DASSLQS	QQTATTPCT
T2/13 DVT(I)	T	AITAAGYSTSYADSVKG	NYNSFDY	AASGLQS	QQTDNNPST
T2/14 DVT(I)	T	NIAADGDTTYYADSVKG	STNAFDY	SASGLQS	QQSSSSPST
T2/15 DVT(I)	T	YINNNGNGTAYADSVKG	ADTCFDY	TASTLQS	QQTYSSPNT

•All clones were selected from libraries based on a single human framework comprising the heavy chain germline genes V3-23/DP-47 and J_H4b and the kappa light chain genes O12/O2/DPK9 and J_K1 with side chain diversity (either NNK (library J) or DVT encoded (library I)) incorporated at 18 different positions in the antigen binding site (underlined). – indicates a deletion of nucleotides.

†This clone has an identical CDRH3 amino acid sequence SSYS as an anti-M scFv (M4) isolated before (9).

Example 7: Parallel selection of antibodies to multiple antigens

SAC is moreover applied to the simultaneous selection of antibodies to multiple antigens. To this end two antibody libraries (I & J) are coselected with three different antigens D, M and T from the same foetal human brain cDNA library (Bussow, *et al.* (1998) *Nucleic Acids Res.* 26, 5007-5008) and analysed the outcome by array screening. Selected (6144) as well as unselected (3072) clones are screened for the presence of cognate interacting pairs, for direct binding to D, M and T and for non-specific binding. After one round of SAC, 92 antibodies are detected against the three antigens. Of these 8 (D), 19 (M) and 4 (T) are confirmed in ELISA.

As before, direct screening for antigen binding allows the detection of specific antibodies not detected by the screen for cognate interaction pairs. Presumably, these arise not only from its greater sensitivity, but also because the direct screen detects potential "false negatives". These may derive from non-cognate phage, which get captured after picking up cognate antigen from the supernatant. In agreement with results from conventional phage selection, binders to antigen D are exclusively isolated from library J, while binders to M or T derived preferentially from library I. Selected binders are specific and do not bind GST, BSA or non-cognate antigens. For selections with library I, few false positives, such as polyreactive antibodies (1/3072) or antibodies specific for the anti- GST IgG (1/3072) or the GST fusion tag (0/3072), are detected. The specific protease cleavage elution, together with the use of just a single round of selection may contribute to the low prevalence of polyreactive binders.

Conversely, library J shows a higher number of polyreactive binders. However, the number of polyreactive clones is very similar before and after SAC selection. Sequence analysis of the anti-D, M and T clones shows that again, SAC yields a highly diverse collection of specific antibodies with all 31 clones being different. Anti-M clones again show Serine-rich VH CDR3 sequences, similar (but not identical) to previously isolated anti-M antibodies (De Wildt, *et al.* (2000) *Nat. Biotechnol.* 18, 989-994), while the anti-T or anti-D antibodies displayed no such similarities (Table 2). 8 of the ELISA-confirmed scFv fragments are purified and affinities measured on BIAcore for both the purified antigens and the corresponding GST-antigen fusion proteins. Affinities are modest for all

antibodies, with dissociation constants (KD) ranging from 0.3 to 1.5μM (Table 3). Affinities for recombinant antigen closely matched affinities for the corresponding GST-fusion proteins.

Table 3: Binding affinities of human scFv fragments isolated using SAC

ScFv fragment	Target antigen	KD (μM) for antigen	KD (μM) for GST-antigen
M2/2	M	0.34	0.39
M2/6	M	1.40	1.54
M2/10	M	0.94	2.10
M3/14	M	1.30	1.61
T2/14	T	1.50	3.50
T2/15	T	0.94	0.69
D2/1	D	0.79	0.46
D2/2	D	0.41	0.28

Example 8: *In vitro* SAC: selection of interactions using avidity capture using *in vitro* display methods.

The method is performed with *in vitro* selection display methods such as ribosome display and puromycin display (Profusion). Libraries of bait and prey ligands are combined on the same transcript. Bait and prey libraries are expressed by *in vitro* transcription translation (ivt). While bait libraries are expressed as soluble fusion proteins comprising the bait ligand fused to a GST hook domain (or any other suitable preferably multimeric fusion partner (unless the bait protein is multimeric in itself)), prey ligand libraries stay attached to the two-cistron transcript either non-covalently via the ribosome ribonucleoprotein complex (Mattheakis *et al.*, 1994 (Proc Natl Acad Sci U S A, 91, 9022-6, Hanes, J. & Pluckthun, A. (1997) *Proc. Natl. Acad. Sci. USA* 94, 4937-4942) or covalently, through a direct puromycin linkage to the transcript (Roberts, R. & Szostak, J. (1997) *Proc Natl Acad Sci USA* 94,

12297-12302). Unless the prey proteins multimeric in itself, prey-gene fusions are also multimerised by N- (or C-) terminal fusion of a suitable multimerising domain, e.g. a leucine zipper peptide. This allows the recovery of interactions with lower affinity (or faster kinetics).

In order to maintain genotype-phenotype linkage in the *in vitro* setting it is desirable to compartmentalise bait-prey library templates prior to ivt, in such a way that on average there is a single bait-prey template per compartment. Ivt then produces a clonal population of bait and prey ligands *in situ* within the compartment. A range of different compartmentation methods are feasible including vesicles (Oberholzer 1999 biochemical and biophysical research communications 261(2): 238-41) water-in-oil emulsions (Tawfik & Griffiths , 1998) or prefabricated nanowells (Borchard et al. 1997, Chemistry and Biology 4(12): 961-8, You et al 1997, Chemistry and Biology 4(12): 969-75). For some applications it may be sufficient to carry out ivt in a diffusion-limiting matrix such as an agarose or alginate gel.

After interactions have been allowed to take place and stable avidity complexes have formed, compartmentalization is broken and avidity complexes and associated bait-prey templates are isolated either by affinity purification via the hook domain or conceivably by a size fractionation step (or combinations thereof). The genotype of selected bait-prey complexes is recovered using PCR. At this stage further rounds of selection can be initiated directly or after recombination (reshuffling) of selected bait and prey genes with the starting (or other suitable) bait and prey libraries. Alternatively, selected bait-prey pairs can be cloned and analysed using established methods, e.g. array screening, ELISA or BiaCore etc.

Example 9: Capture display by SAC

For certain applications it may be desirable to employ an indirect display method, for example when N- (or C-) terminal fusion to another protein interferes with the functionality of the prey protein. SAC allows such indirect display, whereby, for example, a multimeric antibody fragment (e.g. diabody or triabody) expressed solubly in the periplasm is captured by a generic ligand displayed on phage, e.g. protein L, or an anti-tag antibody.

We have shown this for several diabodies displaying a C-terminal cmyc tag and being captured on phage via the anti-cmyc antibody Fab 9E10 as well as for a anti-BSA triabody captured on phage using the generic superantigen ligand protein L (Table 1).

Example 10: Template free SAC: selection and identification of interaction complexes directly at the polypeptide level

This example relates to a method of using the avidity capture concept to analyse polypeptide avidity complexes directly (i.e. without recourse to display methods) using high-resolution MALDI-TOF mass-spectrometry (MS).

As in Example 8, libraries of bait and prey ligands are combined on the same transcript, and expressed by ivt. Both bait and prey libraries would be expressed as soluble fusion proteins comprising the bait or prey ligand preferably (unless the bait or prey proteins are themselves multimeric) fused multimeric hook domains (unless the bait or prey proteins are themselves multimeric). Such hook domains can be GST, leucine zippers etc.

As in Example 8, compartmentation is required to maintain genotype-phenotype linkage. As in Example 8 there is preferably a single bait-prey template per compartment on average.

As in Example 8, after interactions have been allowed to take place and stable avidity complexes have formed, compartmentation is broken and avidity complexes are isolated either by affinity purification via the hook domain, immunoprecipitation or conceivably by a size fractionation step (or combinations thereof). Purified avidity complexes are fractionated further using e.g. gel-electrophoresis and fractions are analysed either directly or after protease digestion using MS.

Example 11: Freeze-frame SAC

This is an alternative method of template free SAC, which allows the capture of short-lived interactions. Potentially it can be applied directly to cellular lysates without the need for genetic library construction and expression.

What is required is an antibody (or other ligand) specific for at least one member of a multicomponent protein complex. This antibody (or ligand) is conjugated to horseradish peroxidase (HRP) (or an alternative source of radicals). When biotin-tyramide is added to the mixture, proteins within a certain radius (e.g. 200Å, but tuneable by reagent concentration) of the radical source are randomly biotinylated. The biotinylated components of the interaction complex can now be cross-linked (e.g. locked in place) by the addition of Streptavidin (or avidin). These cross-linked complexes (freeze-frame complexes) can be fractionated and purified as described above, e.g. by virtue of an anti-Streptavidin antibody and/or an antibody (or superantigen (Protein A, L, or G)) specific for the HRP conjugated antibody. Purified complexes would then be proteolysed and analysed by MS. Since the MW of both Streptavidin and the biotin conjugate are known the covalent biotin modification is unlikely to interfere with peptide analysis.

One the advantages of Freeze-frame SAC is that it lends itself to the investigation of multiprotein complexes involving more than two obligatory interaction partners. For example, in the case of a tri-protein complex ABC, where neither AB, nor AC, nor BC form stable complexes, the methods described above, as well as traditional 2-hybrid methods record a false negative for these bait-prey pairings.

Example 12: Intracellular SAC: selection of intracellular interaction complexes using T7 phage display system

Some proteins cannot be expressed (or are only very poorly expressed) in the periplasm. It is therefore desirable to develop as system for the selection of intracellular interaction complexes. Several such systems exist, including the classic 2-hybrid system (Fields, S. & Song, O. (1989) *Nature* 340, 245), a bacterial 2-hybrid system based on phage lambda repressor (Dove et al 1997 Nature 386(6625): 627-30, Dove et al 1998, Cold Spring Harbor

Symposia On Quantitative Biology 63: 173-80). as well as several split-enzyme approaches (Karimova, *et al.* (1998) *Proc. Natl. Acad. Sci.USA* 95, 5752-5756; Pelletier, *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95, 12141-12146; Mohler et al 1996 Proceedings of the National Academy of Sciences of the United States of America 93(22): 12423-7.

A different approach, based on the principles of SAC (selection by avidity capture (as outlined in this patent)) can be envisioned. This may involve an intracellular phage (e.g. T7, T4 or λ) which allows multiple polypeptide prey fusions to one or several of its capsid proteins (e.g. T7 display system marketed by Novagen) or another intracellular display format, e.g. lac repressor fusions (Cull et al 1992 Proceedings of the National Academy of Sciences of the United States of America 89(5): 1865-9).

Within the same cell multimeric bait proteins are coexpressed (e.g. GST-bait as above) and interaction complexes assemble on the outside of the phage capsid. After cell lysis, phages bearing said interaction complexes can be purified by virtue of the hook protein (GST).

Ideally, genes encoding bait proteins are either encoded as part of the phage genome or on a plasmid bearing appropriate signals (phage origin & packaging signals) to promote copackaging of the plasmid with the phage genome.

All publications mentioned in the above specification, and the publications referenced therein, are herein incorporated by reference. Where the references given are not complete, the complete reference may be easily obtained by consultation of a work such as Medline or *Index Medicus*. Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

CLAIMS

1. A method for selecting first and second molecules according to an interaction between said first and second molecules, comprising:

(i) providing said first and second molecules;

(ii) allowing said first and second molecules to interact, such that a complex comprising each of said first and second molecules in multivalent form is generated; and

(iii) isolating the complex and characterising the constituent molecules.

2. A method for selecting interacting pairs comprising a first molecule and a second molecule which method comprises:

(i) contacting a plurality of first molecules with a second molecule, wherein said second molecule is multivalent and wherein the contacting takes place under conditions that permit interaction of the second molecule with one or more of the first molecules to form a multivalent complex comprising a first molecule and a second molecule;

(ii) isolating a complex formed in step (i); and optionally,

(iii) characterising the first molecule in said complex.

3. A method for selecting interacting pairs comprising a first molecule and a second molecule which method comprises:

(i) contacting a plurality of first molecules with a plurality of second molecules, wherein said second molecules are multivalent and wherein the contacting takes place under conditions that permit interaction of one or more of the second molecules with one or more of the first molecules to form a multivalent complex comprising a first molecule and a second molecule;

(ii) isolating a complex formed in step (i) and;

(iii) characterising the constituent first and second molecules in said complex.

4. A method for selecting first and second polypeptides from first and second pluralities of polypeptides according to an interaction between said first and second polypeptides, comprising:

(i) providing a first library of polynucleotides encoding a plurality of first polypeptides and a second library of polynucleotides encoding a plurality of second polypeptides;

(ii) expressing said first and second libraries to generate the first and second pluralities;

(iii) allowing said first and second pluralities to interact in solution, such that a complex comprising each of said first and second polypeptides in multivalent form is generated;

d) isolating the complex and characterising the constituent first and second polypeptides.

5. A method for selecting binding pairs comprising a first polypeptide and a second polypeptide which method comprises:

(i) expressing a plurality of first polynucleotides to produce a plurality of first polypeptides such that each first polynucleotide which directs expression of a corresponding first polypeptide is associated with said first polypeptide;

(ii) contacting said plurality of first polypeptides with a second polypeptide, wherein said second polypeptide is multivalent and wherein the contacting takes place under conditions that permit multivalent binding of the second polypeptide to one or more of the first polypeptides.

(iii) isolating a complex formed in step(ii); and

(v) isolating a corresponding first polynucleotide that encodes a first polypeptide of said complex.

6. A method for selecting binding pairs comprising a first polypeptide and a second polypeptide which method comprises:

(i) expressing a plurality of first polynucleotides to produce a plurality of first polypeptides such that each first polynucleotide which directs expression of a corresponding first polypeptide is associated with said first polypeptide;

(ii) expressing a plurality of second polynucleotides to produce a plurality of second polypeptides such that each second polynucleotide which directs expression of a corresponding second polypeptide is associated with said second polypeptide;

(iii) contacting said plurality of first polypeptides with a plurality of second polypeptides, wherein said second polypeptides are multivalent and wherein the contacting takes place under conditions that permit multivalent binding of the second polypeptide to one or more of the first polypeptides.

(iv) isolating a complex formed in step(ii); and

(v) isolating a corresponding first and second polynucleotide that encodes a first polypeptide of said complex.

7. A method for selecting binding pairs comprising a first polypeptide and a second polypeptide which method comprises:

(i) expressing a plurality of polynucleotides to produce a plurality of first and second polypeptides such that each polynucleotide which directs expression of a corresponding first and second polypeptide is associated with said first polypeptide;

(ii) contacting said plurality of first polypeptides with a plurality of second polypeptides, wherein said second polypeptides are multivalent and wherein the contacting takes place under conditions that permit multivalent binding of the second polypeptide to one or more of the first polypeptides.

(iii) isolating a complex formed in step(ii); and

(vi) isolating a corresponding polynucleotide that encodes a first and second polypeptide of said complex.

8. A method for selecting binding pairs comprising a first polypeptide and a second polypeptide which method comprises:

(i) expressing a plurality of first polynucleotides to produce a plurality of first polypeptides and a plurality of second polynucleotides to produce a plurality of second polypeptides such that each first and second polynucleotide which directs expression of a corresponding first and second polypeptide is associated within the same entity expressing said first polypeptide

(ii) contacting said plurality of first polypeptides with a plurality of second polypeptides, wherein said second polypeptides are multivalent and wherein the contacting takes place under conditions that permit multivalent binding of the second polypeptide to one or more of the first polypeptides.

(iii) isolating a complex formed in step(ii); an

(iv) isolating a corresponding first and second polynucleotide that encode a first and second polypeptides of said complex.

9. A method according to any one of claims 5 to 8 wherein said second polypeptide is multivalent by virtue of being expressed as a fusion protein to a third polypeptide which multimerises.

10. A method according to any one of claims 5 to 9 wherein said conditions in step (ii) comprise expressing the first polypeptides as a fusion protein to a fourth polypeptide which multimerises.

11. A method according to any one of claims 5 to 9 wherein said conditions in step (ii) comprise expressing the first polypeptides using phage display such that more than one molecule of a first polypeptide is present on the surface of the phage.

12. A method according to any one of claims 5 to 11 wherein each first polynucleotide which directs expression of a corresponding first polypeptide is associated with said first polypeptide by means of phage display.

13. A method according to any one of claims 5 to 11 wherein each first and second polynucleotide which directs expression of a corresponding first and second polypeptide is associated with said first polypeptide by means of phage display.

14. A method according to any one of claims 5 to 12 wherein each nucleotide sequence which directs expression of the corresponding first polypeptide is associated with said first polypeptide by means of compartmentation.

15. A method according to any one of claims 5 to 12 wherein each nucleotide sequence which directs expression of the corresponding first and second polypeptide is associated with said first polypeptide by means of compartmentation

15. A method according to any one of claims 5 to 14 wherein the second polypeptide is present as one of a plurality of second polypeptides expressed from a plurality of second polynucleotides.

16. A method according to any one of claims 5 to 15 wherein step (ii) further comprises contacting said first polypeptides and second polypeptide(s) in the presence or absence of a third molecule or plurality of third molecules, wherein binding of one or more first polypeptides and one or more second polypeptides is modulated by one or more third molecules.

17. A method according to any one of claims 5 to 15 wherein each first polynucleotide is comprised in a separate nucleic acid molecule to the second polynucleotide.

18. A method according to any preceding claim wherein the first and/or second molecule have been modified to increase the valency of said first and/or second molecule.

19. A method according to any preceding claim wherein the first and/or the second molecule comprise at least one chemically reactive group, which group is effective to bond the molecules together covalently upon specific interaction.

20. A method according to any preceding claim, wherein a further molecule covalently cross-links the complex formed between the first molecules and/or the second molecules once assembled.

21. A method according to any preceding claim, wherein a further molecule covalently cross-links the complex formed between the first molecules and/or the second molecules once assembled, e.g. by virtue of a chemical reaction initiated by one or both of the interaction partners.

22. A first molecule or first polypeptide identified by the method of any one of the preceding claims.

23. A second molecule or second polypeptide identified by the method of any one of claims 1 to 21.

24. A third molecule identified by the method of claim 16.

25. Use of a kit comprising a plurality of first molecules and a second molecule or plurality of second molecules in a method according to any one of claims 1 to 3.

26. Use of a kit comprising a plurality of first nucleotides and a second nucleotide or plurality of second nucleotides in a method according to any one of claims 5 to 17.

Figure 1

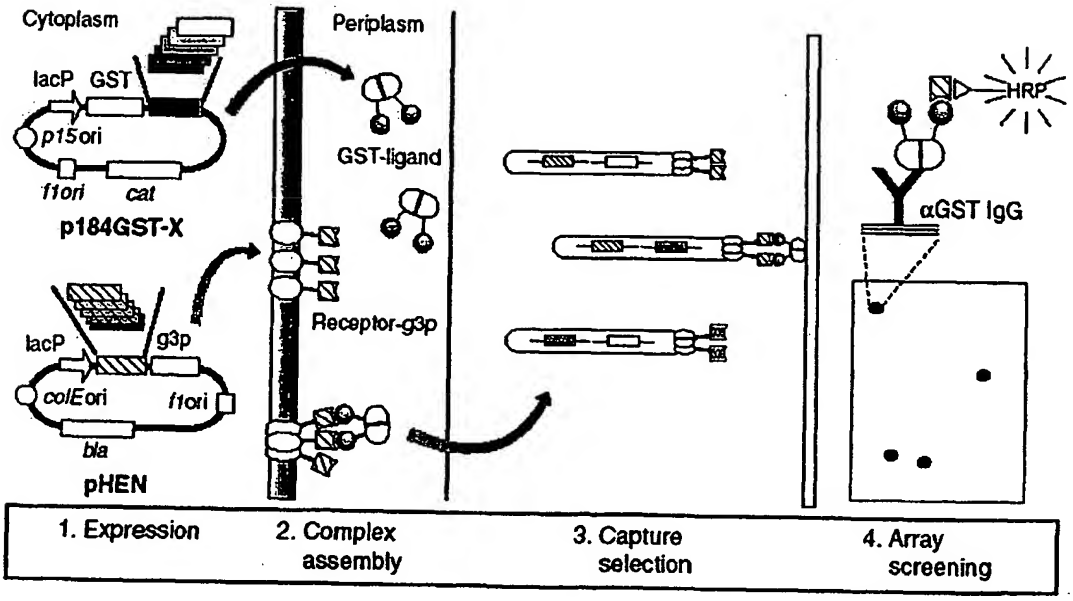
Cytoplasm
Periplasm
lacP
GST
p15ori
f1ori
cat
p184GST-X
GST-ligand
Receptor-g3p
lacP
g3p
colEori
f1ori
bla
pHEN
HRP
αGST IgG
1. Expression
2. Complex assembly
3. Capture selection
4. Array screening

Figure 2

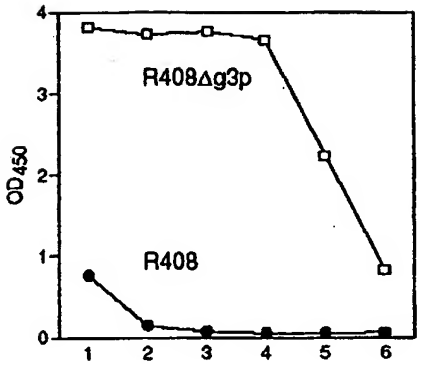

Figure 3

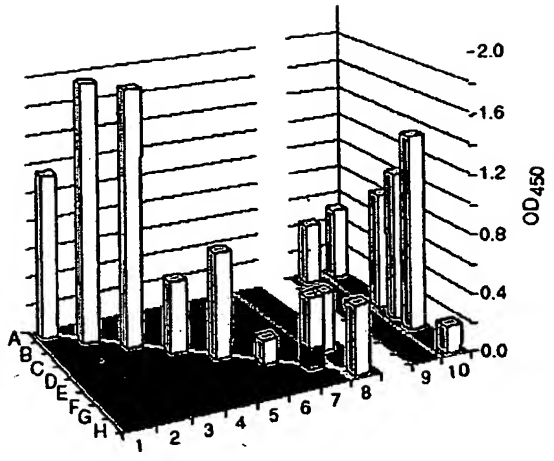

Figure 4

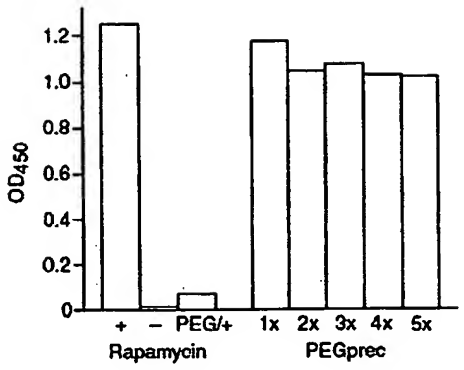

Figure 5